PHYSICS

STEWART

PHYSICS

A TEXTBOOK FOR COLLEGES

BY

OSCAR M. STEWART

PROFESSOR OF PHYSICS, UNIVERSITY OF MISSOURI

GINN AND COMPANY
BOSTON · NEW YORK · CHICAGO · LONDON
ATLANTA · DALLAS · COLUMBUS · SAN FRANCISCO

COPYRIGHT, 1924, BY OSCAR M. STEWART
ALL RIGHTS RESERVED

PRINTED IN THE UNITED STATES OF AMERICA

626.4

𝔗𝔥𝔢 𝔄𝔱𝔥𝔢𝔫𝔞𝔢𝔲𝔪 𝔓𝔯𝔢𝔰𝔰
GINN AND COMPANY · PRO-
PRIETORS · BOSTON · U.S.A.

PREFACE

This text is the result of a long first-hand experience in teaching elementary college physics to those groups of students who have not had special mathematical training. It is designed particularly for those who are taking physics as a part of a general course in arts and sciences, but the author hopes that it will be found comprehensive enough adequately to meet the needs of any elementary college group.

Both the subject matter and the illustrations have been chosen so as to give a large amount of useful knowledge. But there is a more important purpose than merely giving information. The great discoveries and applications of science, to which we are today so indebted, have been due primarily to the development and diffusion of what may be called the modern scientific method of reasoning. The study of physics offers unusual opportunity for acquiring this method. This has been kept in mind throughout the preparation of the text. Too often, however, the student does not get this training because he does not fully grasp the fundamental principles of the subject. It has been the desire of the author to state these principles so clearly and so simply that the student will not only understand them but also learn to use them.

Considerable attention has been paid to the arrangement of material in accordance with the best pedagogical methods. For example, the discussion of any phase of the subject usually begins with a statement of observed facts. Only after the student has had the opportunity of becoming familiar with the facts are deductions or generalizations made. The discussion of a subject is rarely introduced by stating laws or definitions.

The author believes that the only method of gaining an understanding of a principle of physics is by learning a number of applications of that principle. Hence the principles are illustrated by a large number of examples drawn from common experiences. It is hoped that the student will in this way learn to connect the principles of physics with those of daily life.

There are several well-known logical methods of introducing the fundamental concepts of mechanics. In an elementary text it seems desirable to make a choice largely on pedagogical grounds. The *weight* of bodies, with which we are all familiar, has been chosen as a *measure* of force. This choice has made it possible to begin with statics and its application to liquids and gases, and to postpone the more difficult subject of dynamics.

The author has not hesitated to use such terms as *momentum*, *voltage*, and the like before accurate definitions are given. For example, the term *mass*, as something different from *weight*, is introduced before any attempt is made to explain the difference. The student is encouraged to think about that difference and hence to be better prepared when the explanation is finally reached. While this may sometimes appear to give a lack of rigor, yet the advantages are obvious to an experienced teacher. Certainly the student cannot understand definitions or generalizations when he is not familiar with specific facts.

A chapter on meteorology has been introduced, for two reasons: (1) a knowledge of some of the facts and principles is of practical importance; and (2) there are in this subject some first-class examples of the applications of the principles of physics to phenomena which are within the understanding and observation of all.

In the early portions of the text the more familiar units of the British Engineering system are used. Gradually the units of the international scientific system, the centimeter-gram-second system, are introduced; later they are used almost exclusively. The British Absolute system of units, which is rarely found outside of textbooks, is not used.

An attempt has been made to state laws and equations in immediately useful forms. Formulas giving numerical relations between different units are repeatedly introduced in order that the solutions of problems may not be largely exercises in arithmetic.

The ideas and facts of modern physics are freely included; for example, modern notions of the atom and the electron are used wherever helpful. However, some of the more perplexing ideas (such as those of the quantum theory), which properly belong to more advanced texts, are omitted.

The division of the text into a large number of chapters and sections makes its use easier for the instructor who, on account of limited time, is forced to omit portions of the subject matter. The rather large number of divisions has made it possible to segregate some of the more difficult parts of the subject. For example, part of the subject of electrostatics has been reserved for a chapter which is reached only after the student has become more familiar with the concept of potential.

Summaries are given at the end of those chapters where they will be of the greatest service. It is suggested that the student prepare for his own use similar summaries in those cases where they have been omitted.

Answers to only the more difficult problems are given (in an appendix). It is hoped that this plan will tend to prevent the abuse of answers and yet meet the needs of those students who use them properly.

It is expected that this text will be supplemented by suitable demonstration and laboratory experiments. Nothing can take the place of classroom drill as conducted by a good teacher.

It is impossible for the author to give proper credit to all those to whom he is indebted for many of the ideas and methods that have been used. Much help has been received during the actual preparation of the text. The author is under special obligations to Professors H. B. Lemon, of the University of Chicago, and R. A. Millikan, of the California Institute of Technology, who read the manuscript; to Professors S. R. Williams, of Oberlin College, F. R. Watson, of the University of Illinois, and G. W. Stewart, of the University of Iowa, who read the proofs; and to his colleagues Professor H. M. Reese, who read much of the manuscript, and Messrs. H. E. Hammond and R. T. Dufford, who read all the manuscript.

O. M. STEWART

Digitized by the Internet Archive
in 2024

https://archive.org/details/physicstextbookf0000osca

CONTENTS

PART I. MECHANICS AND PROPERTIES OF MATTER

CHAPTER PAGE
I. FORCE 1
II. UNITS OF MEASUREMENT 16
III. LIQUIDS 20
IV. GASES 37
V. UNIFORMLY ACCELERATED MOTION 54
VI. EFFECTS OF FORCES ON MOTION 66
VII. WORK AND ENERGY 90
VIII. SIMPLE MACHINES 109
IX. FRICTION 120
X. VIBRATORY MOTION 129
XI. ROTARY MOTION 139
XII. GRAVITATION 155
XIII. FLUIDS IN MOTION 163
XIV. THE MOLECULAR AND ATOMIC THEORY. SPECIAL PROPERTIES OF SOLIDS AND LIQUIDS DUE TO MOLECULAR ACTION . . 177
XV. ELASTICITY 194

PART II. HEAT

XVI. TEMPERATURE AND ITS MEASUREMENT 203
XVII. EXPANSION 210
XVIII. QUANTITY OF HEAT 228
XIX. WORK AND HEAT 235
XX. THE TRANSFER OF HEAT 250
XXI. CHANGE OF STATE 266
XXII. SOME APPLICATIONS TO METEOROLOGY 293
XXIII. HEAT ENGINES 317

PART III. WAVE-MOTION AND SOUND

CHAPTER PAGE
XXIV. Wave-motion and Sound 325

PART IV. MAGNETISM AND ELECTRICITY

XXV. Magnetism 360
XXVI. Electrostatics 379
XXVII. Simple Properties of Electric Currents 407
XXVIII. Potential Differences; Work and Power in Circuits 416
XXIX. Heating Effects; Joule's Law 435
XXX. Resistance; General Form of Ohm's Law . . . 444
XXXI. Conduction of Electricity through Electrolytes 464
XXXII. Magnetic Fields of Currents; Mechanical Forces on Conductors in Magnetic Fields 477
XXXIII. Induced Electromotive Forces 495
XXXIV. Magnetization of Iron 523
XXXV. Potentials of Charged Conductors; Capacitances of Condensers 531
XXXVI. Electrical Units 550
XXXVII. Conduction of Electricity through Gases . . . 555
XXXVIII. Radioactivity 570
XXXIX. High-Frequency Oscillations and Electrical Waves 580

PART V. LIGHT

XL. Propagation and Photometry 592
XLI. Reflection 606
XLII. Refraction 623
XLIII. Lenses 639
XLIV. Optical Instruments 653
XLV. Spectra and Color Phenomena 665
XLVI. Interference and Diffraction 680
XLVII. Polarization of Light 698
ANSWERS TO PROBLEMS 709
INDEX 715

PHYSICS

PART I. MECHANICS AND PROPERTIES OF MATTER

CHAPTER I

FORCE

Force, 1. Addition of forces, 2. The polygon of forces, 3. Forces in equilibrium, 4. Examples of forces in equilibrium, 5. Vectors, 6. Components of a vector, 7. The sailboat, 8. Forces that produce rotation, 9. The moment of force, or torque, 10. Addition of moments, 11. Conditions for complete equilibrium, 12. The center of gravity of a body, 13. Couple, 14.

1. Force. One of the earliest things we learn in life is that it requires a *force* to lift an object. We soon learn to make estimates of the size or magnitude of the forces required; later, by means of special instruments, such as a lever or a spring balance, we learn to measure forces. Our earliest notions of force are associated with that special type which we call *weight*. This, we notice, always acts in a vertical direction. We soon learn, however, that forces may act in any direction, and that if we wish to accomplish something which requires a force we must pay as much attention to the direction of the force as to its magnitude. In general, it may be said that a force is completely described, or specified, only when both its *magnitude* and *direction* are given.

In measuring the magnitude of forces various units are used. In this book we shall generally use the *pound* and the *dyne*. The first is commonly used in English-speaking countries; the second is the international scientific unit. The dyne is an extremely small unit:

$$1 \text{ lb.} = 444{,}823 \text{ dynes.}$$

In this chapter will be given some of the methods and principles which must be known before a more detailed study of the relation of forces to other physical quantities is undertaken.

2. Addition of forces. When forces act in the same direction, they are added arithmetically. If to a weight of 10 pounds lying on a platform we add 15, and then 20 more, the total load will be 45 pounds. These forces can be added in this way only because they are in the same direction—they all act vertically downward. Forces which act in the same direction or in opposite directions are spoken of as *parallel* forces. In all cases parallel forces are added algebraically: those acting in one direction are called positive, and those in the opposite direction negative. Suppose that a given body has a number of forces acting on it—forces of 15, 25, and 30 pounds in one direction, and forces of 20 and 10 pounds in the opposite direction. These forces can be added as follows:

$$\begin{array}{r} +15 \\ +25 \\ +30 \\ -20 \\ \underline{-10} \\ +40 \end{array}$$

Hence the resultant force is one of 40 pounds acting in the direction called positive.

When the forces are not parallel, ordinary arithmetical or algebraic methods cannot be used. For example, suppose that one wished to find the sum, or resultant, of two forces making an angle of 60 degrees with each other—one of 50 pounds, and the other of 20 pounds. A method available for solving this is a graphical method, called the *parallelogram of forces*. In Fig. 1 the two forces are represented in direction and magnitude by the two lines *A* and *B*. The lengths of these lines are drawn to *scale*; that is, the line *A* is made twice as long as the line *C*, and *B* is made five times as long as *C*. The length of *C* represents 10 pounds. The length of the line *C* is entirely optional: any convenient length may be used. The parallelogram is com-

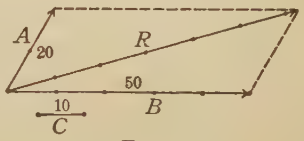


FIG. 1

pleted by drawing the broken lines. The diagonal R, both in direction and magnitude, represents the sum, or resultant, of the two forces. The magnitude of R is determined by comparing its length with the length of the line C. In this way it is found that R represents a force of about 62 pounds. The diagonal R is sometimes called the *geometrical sum* of A and B.

In adding forces by the use of the parallelogram method two simple rules must be observed: (1) the two forces and the resultant are drawn from the same point, sometimes called the *origin*; (2) the arrows, which indicate the direction of the forces, should all point away from the starting-point, none toward it.

The parallelogram method may be used when there are more than two forces to be added. First, the resultant of any two of the forces is found by drawing a parallelogram. To this resultant the third force is added by the same method, a new resultant being found. This process is followed until all the forces have been added. But when more than two forces are involved, the work is cumbersome. An improved method, called the *polygon of forces*, is explained in the following section.

3. The polygon of forces. It will be noticed by referring to Fig. 1 that it is not necessary to draw both the broken lines to determine the length and direction of R. All that is necessary is to draw two sides of the parallelogram. This is clearly seen in Fig. 2. Here B is drawn at the upper end of A, instead of at the lower end as in Fig. 1. The advantage of this method is more striking when one tries to find the resultant of several forces.

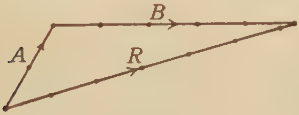


FIG. 2

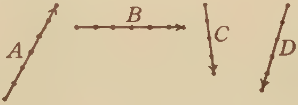


FIG. 3

Let us try to add the four forces $A = 6$ pounds, $B = 6$ pounds, $C = 4$ pounds, and $D = 5$ pounds. These forces are drawn to scale and in the proper directions in Fig. 3. In Fig. 4 is shown the method of adding these four forces. The line A is drawn first. To the end of A is attached B, care being taken that B has the proper length and direction. To the end of B is added C, and to that D. The line R, connecting the origin O and the end of D, is the resultant force. From the length, R is found to be about 8.5 pounds.

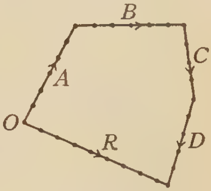


FIG. 4

It should be noted that in Figs. 2 and 4 the lines are so connected that two arrowheads do not point to any one junction, or corner, except in the case of the resultant. When connected as shown in Fig. 5, incorrect results will be obtained. The student should satisfy himself of this by an actual trial of simple cases, such as the forces of Fig. 2.

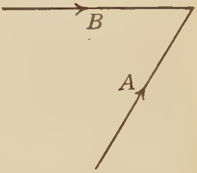


FIG. 5

4. Forces in equilibrium. Whenever a number of forces acting at the same place have a resultant which is zero, the forces are said to be in *equilibrium*. The lines which represent forces in equilibrium will form a *closed* figure when drawn by the method explained in the last section. For example, the forces represented in Fig. 6 are in equilibrium; their sum is zero. That they are in equilibrium may be proved as follows: The resultant of the forces AB and BC is a force AC acting from A toward C. If a third force equal and opposite to this resultant is added, the sum will be zero, and hence equilibrium will be obtained. But such a force has been included. It is the force CA, acting from C toward A. Hence the three forces AB, BC, and CA, which form the closed triangle ABC, are in equilibrium.

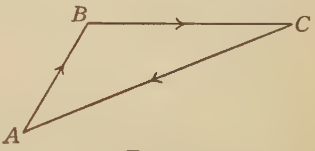


FIG. 6

In order that lines forming a closed figure should represent forces in equilibrium, it is necessary that the forces act in the same sense around the figure. For example, the directions A to B, B to C, and C to A, in Fig. 6, are all in the same sense. In Fig. 4 the arrowheads do not all point in the same sense; but if the resultant R were replaced by a fifth force equal and opposite to R, the heads would all point in the same sense. Then the five forces would be in equilibrium, for they would meet the two conditions: (1) that the lines representing these forces form a closed figure, (2) that the arrowheads point in the same sense around the figure.

Whenever the resultant of a system of forces acting at a point is zero, lines representing these forces will form a closed polygon, with the direction of the forces in the same sense around the figure.

5. Examples of forces in equilibrium. A rope (Fig. 7) has a weight W in the middle. The problem can be stated thus: Given the inclination of the two parts of the rope, and the weight W of the load, to find the tension, or force, in the two parts of the rope. It is a matter of experience that when a body is at rest, the resultant of all the forces acting on it is zero. Hence the forces acting on A, shown in Fig. 8, are in equilibrium and, when properly placed together, will form a closed triangle (Fig. 9). In this problem the magnitude of W is known; hence that line is drawn first, care being taken to get its direction and length correct. The magnitudes of F_1 and F_2 are not known, but the directions are. Hence the lines F_1 and F_2 are drawn from the ends of the line W in the correct directions until they intersect. Then the lengths are carefully measured, and the magnitudes of the forces F_1 and F_2 determined from the scale of the figure.

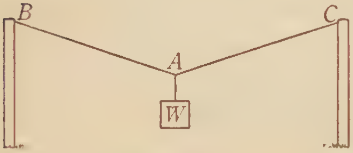


FIG. 7

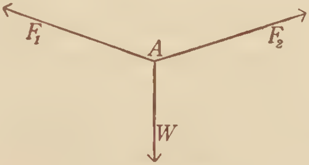


FIG. 8

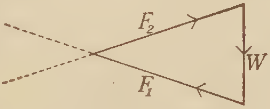


FIG. 9

When the angle BAC of Fig. 7 is 120 degrees, the triangle of Fig. 9 is equilateral, and the forces F_1, F_2, and W are all equal. If the angle is nearly 180 degrees, a small load W will produce a large tension on the ropes AB and AC. This is the explanation of the great tension produced in taut wires when they become loaded with ice during an ice storm. The greater the slack in the wires, the less will be the tension. The hammock is an example of this problem. As is well known, there is less danger that the hammock ropes will break if they are more nearly vertical.

In the crane shown in Fig. 10 a weight W is supported in its position by forces supplied by the rope AB and the boom AO. These forces act at the point A: the rope, with a force F_1, in the direction AB; and the boom, with a force F_2, in the direction parallel to its length.* Suppose that the weight W and the inclinations of the boom and the rope AB are known, and that it is desired to find the forces exerted by the boom and the rope AB. Since the point A is at rest, the forces acting on it must be in equilibrium. Hence the lines representing the forces form a closed figure. Since the magnitude and direction of W are known, the line W of Fig. 11 is properly drawn as regards direction and length. Starting at D, a line DE is next drawn parallel to the force F_1. Next, from C is drawn a line CF parallel to the boom. These lines intersect at a point H.

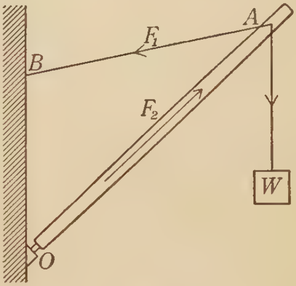


FIG. 10

Since the lines representing the three forces form a closed figure, the lines DH and HC must represent in magnitude and direction the forces F_1 and F_2. The magnitudes of F_1 and F_2 are determined by measurements on the diagram.

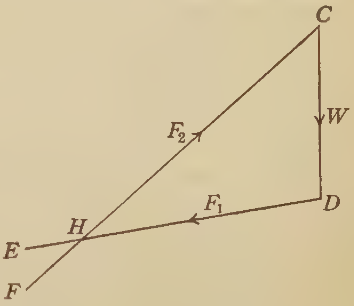


FIG. 11

The problem given above may be changed in several different ways. It may be that the forces are all known, but the directions of F_1 and F_2 unknown. In that case the line W is drawn first, and the problem is solved by finding the triangle which can be constructed by adding two lines of known length. As this is a familiar geometrical problem, the details need not be given here.

---

* The weight of the boom is neglected.

Or, the direction and magnitude of the forces F_1 and F_2 being given, it may be desired to find the weight W which can be supported. The student should be able to explain the necessary steps.

6. Vectors. Lines which by their direction and magnitude represent forces are called vectors. A vector may be used to represent not only force but *any quantity which has direction and magnitude*. Force, velocity, acceleration, and momentum are vector quantities. Those quantities which have magnitude but do not have direction are called *scalar* quantities. The number of acres in a farm, the balance in a bank account, the population in the state, a baseball score, and most of the quantities with which we deal in everyday life are scalar quantities. *Scalars are added by ordinary arithmetical methods, but vectors are added by geometrical methods*. The methods which have just been given for adding forces can be used for any kind of vector quantities. Figs. 9 and 11 are special cases of vector diagrams.

7. Components of a vector. It often happens that not all of an applied force is effective. In Fig. 12 a force F acting in the direction AB is used to pull a barrel up an incline. Only part of this force is effective in rolling the barrel up the incline. In the figure the line AC is drawn parallel to the surface, and the line CB perpendicular to AC. If we add the two vectors AC and CB, the resultant is the vector AB, or F. The vectors AC and CB are called *components* of the vector AB. The vector AC represents that part of F which produces the motion. The other part, CB, exerts a force perpendicular to the surface of the incline and hence is not effective in producing the desired motion.

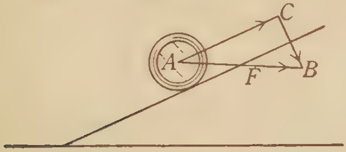


FIG. 12

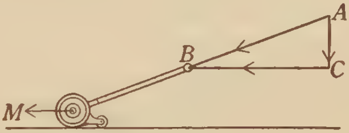


FIG. 13

In Fig. 13 the line AB represents a force used to push a lawn mower in the direction M. The force AB may be broken up into two components AC and CB. The component CB produces the motion in the direction M, but the component AC pushes the mower against the ground and does not assist in moving it.

A sled coasting downhill is acted upon by the weight of the sled and its occupants. In Fig. 14 this weight is represented by the vector AB. The right triangle ABC is constructed. The side AC is that component of the weight AB which is effective in moving the sled down the hill. The component CB is a force perpendicular to the ground and contributes nothing to the motion. It is obvious from the figure that if the hill is steeper, the effective component AC is larger.

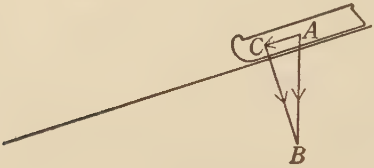


Fig. 14

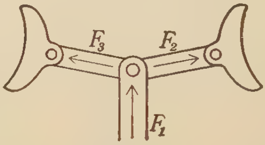


Fig. 15

In a form of lever called the knee, or toggle, joint a force applied at F_1 produces larger forces, F_2 and F_3 (see a simplified form in Fig. 15). This lever is often utilized in the construction of brakes, such as those used on locomotives. In the vector diagram of Fig. 16 it is shown that F_2 and F_3 are the components of F_1, for the force F_1 is the vector sum of the two forces F_2 and F_3.

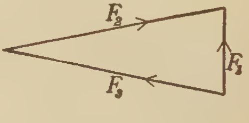


Fig. 16

Any vector may be broken up into two or more components, the only condition being that the sum of these components must equal the original vector.

8. The sailboat. The explanation of how a sailboat can beat into the wind forms an interesting application of the resolution of vectors into components. In Fig. 17 let B represent the direction of the motion of the boat, and W the direction of the wind. The wind acting on the sail can exert a force on the sail only in a direction perpendicular to it. In the figure the vector AD represents the force with which the wind is acting on the sail SS'. The force AD is transmitted through the mast and rigging to the boat. In the lower part of the figure the force AD is shown broken up into the two components *AC* and *CD*. *CD* is the force which tends to make the boat move sidewise. But a sailboat is furnished with a centerboard, or keel, which tends to prevent side drift of the boat. The component *AC* is the effective one, the force which causes the boat to travel in the direction *B*. Usually there is some side drift, or leeway, so that the motion is not exactly in the direction *B*.

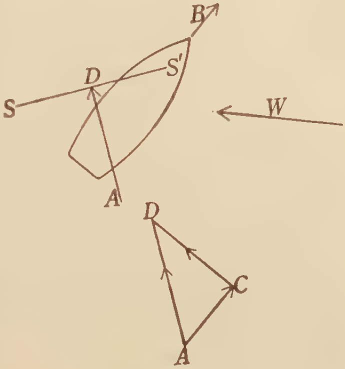


FIG. 17

9. Forces that produce rotation. So far we have been dealing with forces which tend to produce linear motion of the bodies on which they act. But this is not the only type of motion. For example, in the case represented by Fig. 18 the body *A* has two forces acting on it which are equal and opposite in direction. From what has been stated in the preceding sections, one might assume that the sum of these two forces is zero and that no motion would be produced. This is not a correct interpretation of the principles we have had. The two forces will annul each other only when they lie in the same straight line. In the case shown in Fig. 18 a rotation of the body will be produced.

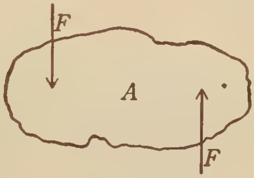


FIG. 18

It now remains to find out how to treat forces that tend to produce rotation.

10. The moment of force, or torque. In producing rotation about any axis, we learn from experience that a force increases in effectiveness when we apply it farther from the axis. In opening a gate or a door we usually apply a force as far from the hinges as we can. A force applied at a considerable distance from the axis is said to have a greater importance, or moment, than the same force applied nearer the axis. To measure this importance, or moment, *the product of the force and the distance of the force from the axis* is used. This product is called the *moment of force* or, frequently, the *torque*.

The precise meaning of the expression "distance of the force from the axis" must be explained. To find this distance a line is drawn from the axis perpendicular to the line of the direction of the force. In Fig. 19 the lengths of the lines a and b are the desired distances. The moment of the force F_1 is F_1a, and that of F_2 is F_2b.

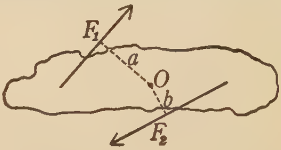


FIG. 19

In the case of the lever shown in Fig. 20 the force F applied at one end will balance a larger force, a weight W, at the other end. The moment of force produced by the force F is Fb, where b is the distance to the axis or fulcrum. The greater the distance b, the greater the moment of the force produced by the force F. The moment due to the weight—that is, the product Wa—tends to turn the lever in a direction opposite to that in which the moment Fb tends to turn it. When these two moments are equal and opposite, the lever is balanced, or in equilibrium. When the two moments of Fig. 20 are equal, we have

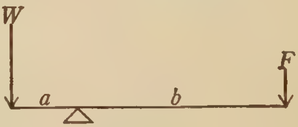


FIG. 20

$$Fb = Wa,$$

or

$$F = \frac{a}{b} W. \tag{1}$$

This equation states the *law of the lever.*

11. Addition of moments. When two moments of force tend to produce rotation in the same direction, the resultant moment is the sum of the two; when the two act in opposite directions, the resultant is the difference between them. It is customary to call those moments tending to produce rotation in a counterclockwise direction *positive*, and those tending to produce rotation in a clockwise direction *negative*. A simple rule is at once formed: *The resultant of several moments is their algebraic sum.** This gives not only the magnitude of the resultant but also its direction, for the algebraic sign gives the direction. The resultant of the moments about the axis at O (Fig. 21) is obtained by adding as follows:

---

*This discussion applies only to forces that lie in the same plane. For the more general case, methods explained in Chapter XI must be used.

$$\begin{array}{rr} (10 \times 5) = & 50 \\ -(15 \times 3) = & -45 \\ -(25 \times 2) = & -50 \\ -(20 \times 4) = & -80 \\ \hline \textit{Resultant moment} = & -125 \end{array}$$

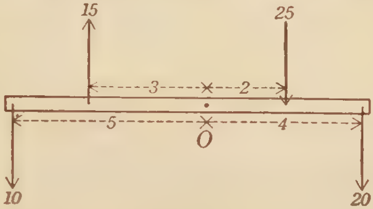


Fig. 21

The resultant moment is clockwise and equal to − 125. This is equivalent to a force 25 applied at a distance 5, or a force 5 at a distance 25, or a force 12.5 at a distance 10, and so on.

12. Conditions for complete equilibrium. It has been pointed out that if a number of forces act on a body which is at rest, they will not produce a translational motion if the vector sum of the forces is zero. Similarly, no rotational motion will be produced if the sum of the moments of force is zero. These are called the *first and second conditions of equilibrium*. To summarize:

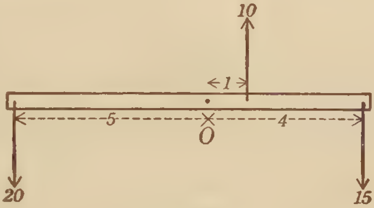


Fig. 22

The first condition of equilibrium is that the vector sum of all the forces acting on the body must be zero.

The second condition of equilibrium is that the sum of the moments of force about any axis in the body must be zero. (This is sometimes called the *law of moments.*)

As an illustration of the use of the two conditions of equilibrium, the magnitude of the force and the place to apply it to produce equilibrium in the case shown in Fig. 22 will be computed.

The first condition tells us that the total downward force must equal the total upward force. Since the total downward force is 35 pounds and the upward force is only 10, *there must be added an upward force of 25 pounds.* The location of the point where this force must be applied can be determined by the second condition.

The second condition demands that the sum of the moments of force must be zero. The moments of the given forces are

$$\begin{aligned} (20 \times 5) &= +100 \\ (10 \times 1) &= +\ \ 10 \\ -(15 \times 4) &= -\ \ \underline{60} \\ \mathit{Sum} &= +\ \ 50 \end{aligned}$$

Hence there must be added a moment of – 50 to produce equilibrium. But we have already seen that the force must be 25 pounds and must be upward. In order that this force may have a moment of – 50, it must be added at a distance of 2 to the left of the axis.

In solving such a problem as this, it does not make any difference what point is chosen as the axis. When equilibrium is produced, the body has no tendency to rotate about an axis placed at any point. In applying the second condition we can therefore choose the center or end or any convenient point as the axis.

13. The center of gravity of a body. In problems such as the one discussed in the last section the weight of the body may produce an appreciable effect. In crowbars and many other forms of levers the weight usually cannot be neglected. In order to include the weight, we must be able to compute the moment of force that it produces. The weight of a body is the resultant of all gravitational forces acting on the different parts of the body. The resultant of these downward forces passes through a point called the *center of gravity*. If no other forces are applied, the body will be in equilibrium when an upward force equal to its weight is applied at its center of gravity. Hence one may assume, in computing the moment due to the weight, that all the weight acts at the center of gravity.

Usually it is not difficult to find the center of gravity by experiment. If a bar is balanced on a knife edge, the knife edge must be directly under the center of gravity. Or when a body is hung by a string, the imaginary line which is a continuation of the stretched string will pass through the center of gravity. If the body is now supported by the string attached at a different point of the body, another imaginary line which passes through the center of gravity is obtained. The intersection of these two lines gives the position of the center of gravity.

As an example take the case of a plank, weighing 50 pounds, balanced on a fence (Fig. 23). On one end of the plank is a girl weighing 30 pounds, on the other a boy weighing 60 pounds. How far from the girl must the fence be if the plank is 12 feet long? Let x be the distance of the girl from the fence. The moment of force due to the girl is $+30x$. The center of gravity of the plank is 6 feet from the girl, or a distance of $x-6$ feet from the fence. The moment due to the weight of the plank is $+50(x-6)$. The boy is at a distance of $12-x$ from the fence. The moment he exerts will be $-60(12-x)$. The sum of these moments must be zero. Hence

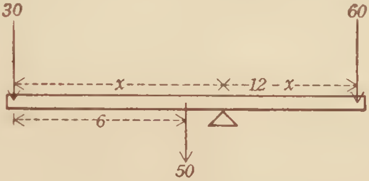


FIG. 23

$$30x+50(x-6)-60(12-x)=0.$$

From this $$x=7.29 \text{ ft.}$$

14. Couple. Cases are often found where two equal forces are acting in opposite directions but not in the same straight line. When one winds a clock, he uses two equal and opposite forces to turn the key, and the forces are not in the same straight line. This particular combination of forces is called a *couple*.

The total moment of force produced by a couple can be computed as follows: Let one of the forces be at a distance x from some point (Fig. 24); then its moment about this point is Fx. If the other force is at a distance y from this same point, its moment is Fy. The resultant moment is

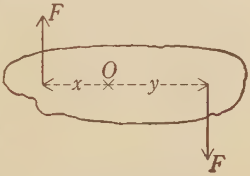


FIG. 24

$$Fx+Fy,$$

or $$F(x+y).$$

But $(x+y)$ is the distance between the two equal forces. Hence *the moment of a couple is the product of one of the forces by the* *distance between the forces.* This distance must be the shortest distance; that is, the perpendicular distance between the two forces.

A single force will not balance a couple. To produce equilibrium, a couple equal and opposite to the existing couple must be applied.

PROBLEMS

1. A man travels 4 mi. south, 3 mi. west, 6 mi. north, and 5 mi. east. Find his distance and direction from the starting-point.

2. A boat is traveling with a velocity of 4 mi./hr. in a direction at right angles to the current. If the current is 3 mi./hr., find by a vector diagram the direction and magnitude of the resultant velocity.

3. Find graphically the resultant of the following three forces: (*a*) 50 lb. north, (*b*) 100 lb. east, and (*c*) 75 lb. northeast.

4. Two forces, 100 lb. and 200 lb., are acting at right angles to each other. Find by a graphical method the magnitude and direction of the resultant force.

5. Find the vector sum of the following displacements: east 4 mi., north 2 mi., east 1 mi., and south $\frac{1}{2}$ mi.

6. Add the forces *A*, *B*, *C*, and *D* of Fig. 3, taking them in a different order from that in Fig. 4. Does the order in which the forces are taken make any difference in the result?

7. A weighted brush with a long handle is used for polishing floors. In which case is there greater pressure on the floor, when pushing the brush or when pulling it? Explain, using a diagram.

8. Given a force *OC* of 60 lb. Show by a diagram how to find two forces, *AB* and *CD*, at right angles to each other, that will just balance the force *OC*.

9. A force of 50 lb. acts at an angle of 60° with the horizontal. Show how to find graphically the horizontal and vertical components.

10. The resultant of two forces acting at right angles is 65 lb. If one of the forces is 7 lb., find the value of the other.

11. A 200-pound weight is supported by two ropes, each making an angle of 30° with the horizon. Find the tension of each rope.

12. A car on a level track is pulled by a force of 100 lb. in a direction making 45° with the track. What is the component which is effective in moving the car?

13. Weights of 3 lb. and 12 lb. balance on a lever the longer arm of which is 2 ft. long. What is the length of the shorter arm if the weight of the lever can be neglected?

14. A uniform bar 60 in. long weighs 20 lb. If a load of 150 lb. is supported at one end, how much force must be supplied to the other end in order that there may be equilibrium about a support placed 10 in. from the loaded end?

15. A wheel of an automobile is lifted off the ground by a lever 8 ft. long. If the load of 800 lb. is applied at one end at a distance of 6 in. from the fulcrum, how much force must be applied at the other end of the bar? The bar, which is uniform in shape, weighs 20 lb.

16. A uniform bar is 60 in. long and weighs 5 lb. If a weight of 6 lb. is suspended from one end and a weight of 10 lb. from the other, at what point must the bar be supported in order that it will balance?

17. A bar *AB*, 10 ft. long, balances at a point 4 ft. from the end *A*. When a weight of 20 lb. is placed at *A* and a weight of 40 lb. at *B*, the bar balances at a point 6 ft. from *A*. Find the weight of the bar.

18. In order to lift a 500-pound weight placed at one end of a 15-foot bar, a force of 100 lb. is applied at the other end. If the bar is of uniform size and weighs 25 lb., where must the fulcrum be placed?

19. The boom of a crane makes an angle of 45° with the vertical. A tie rope which is at right angles to the boom is fastened at the outer end of the boom. If a 600-pound weight is hung at the outer end of the boom, find (*a*) the tension in the tie rope and (*b*) the thrust of the boom.

20. Two men are carrying a load of 150 lb. on a uniform bar 10 ft. long. Where must the load be placed in order that one of the men may carry three fourths of the total load? The weight of the bar is 50 lb.

CHAPTER II

UNITS OF MEASUREMENT

Fundamental units, 15. Derived units, 16. Units of length, 17. Units of mass, 18. Units of time, 19. Units of force, 20.

In any kind of quantitative work it is a great advantage to have a convenient system of units. The common British system of weights and measures is a glaring example of what not to have. In this system there are a great many different units for measuring the same thing, and, what is worse, there is often no simple way of changing from one to another. For example, to measure volume we may use cubic inches, cubic feet, cubic yards, bushels, barrels, gallons, quarts, pecks, etc. Of some of these there are two kinds; for example, the wet gallon and the dry gallon. The difficulty sometimes encountered of changing from one of these to another is shown by the work necessary to change cubic feet to bushels, a not uncommon computation. The use of different units for measuring the same quantity is not objectionable provided the relationship of the units is simple. No one has trouble in reducing dollars to cents, owing to the simplicity of the decimal system. This system is so satisfactory that it is used the world over in scientific units.

15. Fundamental units. In any system it is necessary at the beginning to assume certain units arbitrarily. In each of the commonly used systems there are three of these *fundamental* units. The fundamental quantities in the international scientific system are *length, mass*, and *time*. The units of these are, respectively, the *centimeter*, the *gram*, and the *second*. The initial letters of the units give the name to the system, which is known as the C.G.S. system. In the British Engineering (B.E.) system the fundamental quantities are *length*, *force*, and *time*. The units of these are the *foot*, the *pound*, and the *second*. It should be noticed that while *mass* is a fundamental quantity in the C.G.S. system, *force* is the fundamental quantity in the British Engineering system. The difference between mass and force, and the important reasons for making a distinction, will be discussed later.

The C.G.S. system is used in scientific work in all countries. Engineers in English-speaking countries generally use the British Engineering system, although the use of metric units is constantly increasing. The electrical and magnetic units used by engineers are based on C.G.S. units.

16. Derived units. Frequently the unit of a quantity is obtained from other units. For example, the cubic foot is a unit of volume obtained from the unit of length; a foot-pound is a unit of work derived from the units of length and force. Such units are called *derived* units. A liter is a volume equal to 1000 cubic centimeters. It is a derived unit based on the unit of length—the centimeter. In the C.G.S. system nearly all the units are derived. In general it is simpler to use *derived* units than *arbitrary* ones. For example, if one knows the dimensions of a bin in feet, it is easier to compute the volume in cubic feet, a derived unit, than in bushels, an arbitrary unit. An arbitrary unit is one chosen without regard to the other existing units.

17. Units of length. The international meter was defined as the distance between the ends of a platinum bar, the *mètre des Archives*, made by Borda and preserved in the national archives of France. Between 1875 and 1887 the International Bureau of Weights and Measures, with headquarters in a suburb of Paris, made thirty-one copies of the Borda meter. Each of these was made of an alloy of 90 per cent platinum and 10 per cent iridium. In each case the meter was the distance between two fine parallel lines ruled on the surface of the bar. The one which agreed most closely with Borda's meter was declared to be the International Prototype Meter. The others were distributed among the governments which participated in the support of the Bureau. The United States secured two of these, Nos. 21 and 27. They are preserved in the Bureau of Standards in Washington.

The one-hundredth part of the meter is the centimeter, the basic unit of length in the C.G.S. system. For small distances the millimeter (one tenth of the centimeter) is often used, while for large distances the kilometer (1000 meters) is used. In special cases other units of length are used in scientific work. For example, in measuring microscopic distances the *micron* (one millionth of a meter) is used. But in all cases the unit used should be some decimal multiple of the meter.

In the United States the meter is the legal standard for both metric and common units. The yard was defined in terms of the meter by an act of Congress, July 28, 1866:

$$1 \text{ yd. (U.S.)} = \tfrac{3600}{3937} \text{ m.}$$

From this,

$$1 \text{ m.} = 39.37 \text{ in.}$$

The British imperial yard is defined in a different way, but for nearly all purposes the United States and British yards may be regarded as equal.

18. Units of mass. The *gram* is approximately the mass of 1 cubic centimeter of water at 4° C. The precise value of the gram is defined as being the one-thousandth part of the standard kilogram, a mass carefully preserved by the International Bureau of Weights and Measures. The United States government obtained two accurate copies of the standard kilogram about the time the two meters were delivered.

19. Units of time. There are two different time systems used. The basis of one of these is the time of revolution of the earth on its axis. This is called the sidereal day, and 1/86400 of it is called the *sidereal second*. The more common system is based on the apparent motion of the sun. When the sun is on the meridian, it is said to be *solar noon*. The time between two successive solar noons is the *solar day*. As solar days at different seasons are not of the same length, the *mean solar day* is used, and the *mean solar second* is 1/86400 of this. The mean solar second is about $1\frac{1}{365}$ sidereal seconds. In all the various systems of scientific units the mean solar second is used as the unit of time.

20. Units of force. The *pound* in the United States is legally defined as the weight of 1/2.204622 *kilogram*. The weight of this mass varies slightly from place to place on the earth's surface. For this reason it is customary in very accurate work to use as the standard the weight at a place where a falling body would gain in one second a velocity of 32.1739 feet per second. This is approximately the velocity acquired at sea level at 45 degrees latitude.

In the C.G.S. system the unit of force is the *dyne*, a derived unit. It will be defined later.

The weight of various masses can be used as force units. Sometimes the weight of a kilogram is used as a unit of force. But this unit should never be called a kilogram. It is better to say the weight of a kilogram, or a kilogram-weight. The kilogram is a mass unit (not a force unit) in the systems used in this country.

CHAPTER III

LIQUIDS

Three forms of matter, 21. Pressure and force, 22. Some facts regarding pressures in liquids, 23. Derivation of the pressure-depth formulas, 24. The hydrostatic paradox, 25. Units for measuring pressure, 26. Transmission of pressure; Pascal's principle, 27. Archimedes' principle, 28. Theoretical proof of Archimedes' principle, 29. Examples of Archimedes' principle, 30. Specific gravity; density, 31. Measurement of the specific gravity of solids, 32. The specific gravity of liquids, 33. The compressibility of liquids, 34.

21. Three forms of matter. Matter exists in three forms: *solids*, *liquids*, and *gases*. Sometimes the last two of these are grouped under the term *fluids*. The property possessed by solids that differentiates them from liquids and gases is called *rigidity*, the tendency to maintain shape. There is no hard and fast line of separation between solids and liquids, for some solids if left alone long enough exhibit properties of liquids; thus pitch and some of the waxes will flow if given time enough.

Many substances change freely from one form to another. We are familiar with water in the three forms, solid, liquid, and gaseous, but it is not so commonly known that many other substances can exist in the different forms. Metals, such as iron, copper, and lead, when raised to sufficiently high temperatures become gases. On the other hand, all known gases have been changed into liquids at low temperatures and, with the exception of helium, have been solidified at still lower temperatures.

In reading over the discussion of the familiar properties of liquids and gases the student should avoid confusing a knowledge of the facts with an understanding of the fundamental principles involved. In scientific work a knowledge of the facts is only the first step. When the principles are understood and the ability to use them is acquired, it will be found not only that it is easier to remember the facts but that invention and the prediction of additional facts become possible.

22. Pressure and force. It is necessary to distinguish between two concepts, *pressure* and *force*. In the apparatus illustrated in Fig. 25, water exerts forces on two pistons. The *pressure* on each of these is numerically equal to *the force exerted on a unit area*. If the force is expressed in pounds and the area in square inches, pressure is the number of pounds acting on each square inch of the piston.

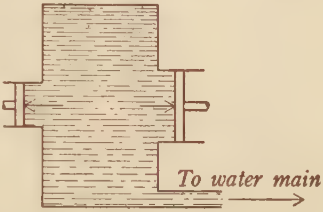


FIG. 25

An experiment made with this apparatus would show that the forces on the two pistons are not the same. From the numerical values obtained one could verify, within experimental errors,

$$\frac{\text{Force on large piston}}{\text{Area of large piston}} = \frac{\text{force on small piston}}{\text{area of small piston}}.$$

But each of these fractions is equal to the *pressure*. Hence the result can be stated in either of two ways:

1. The force acting on a piston is proportional to its area.
2. The pressures on the two pistons are equal.

The second statement is simpler than the first. This is an example of what is generally true: that it is simpler to state facts and laws about fluids in terms of pressures than in terms of forces.

23. Some facts regarding pressures in liquids. Simple experiments show that

1. There is a pressure not only at the bottom but everywhere throughout a liquid.
2. Pressure increases with depth.
3. At any place in a liquid there is pressure in every direction.
4. At any place in a liquid the pressure is equal in all directions.
5. Pressure in a liquid at rest always acts perpendicularly to a surface.
6. Pressure in a liquid at rest is the same at all points on the same level.

One of the simplest ways of verifying most of these facts is by the use of the gauge illustrated in Fig. 26. A piece of thin sheet rubber is tied over the large open end of a thistle tube. The tube is cut off near the bulb, and a piece of rubber tubing is slipped over this shortened end. The other end of the rubber tubing is connected to a U-shaped glass tube partly filled with water colored with a little ink. When the thistle tube is immersed in water, pressure on the membrane is shown by the displacement of the colored water in the U-tube. The bulb with the membrane can be raised or lowered, turned sidewise, or inverted. The student should try to imagine what must be done with the gauge to verify facts 1, 2, 3, 4, and 6.

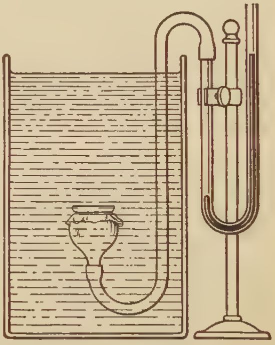


Fig. 26

There are many well-known phenomena which when recalled serve to verify these facts. All who read the newspapers know that divers and submarine boats cannot descend to great depths on account of the enormous pressures at those depths. We are familiar with the disastrous results produced when the side pressure on a dam breaks it. There are abundant evidences of the existence of an upward pressure in a liquid. When a stick is forced end down into water, it is the upward pressure on the end that pushes the stick out. If a card is held over the smooth end of a glass tube while the tube is lowered into the water, the water pressure will hold the card tightly up against the tube.

That there is an upward pressure in the interior of a liquid should be expected as a result of very simple reasoning. Imagine a piece of thin tissue paper submerged in a liquid. If the paper is lying in a horizontal position, the weight of the liquid above it is pressing down on it. Since the paper is frail and not capable of sustaining a load, one is forced to the conclusion that there must be an upward pressure on the underside of the paper. If the paper is removed, these forces still exist. In a similar way it may be seen that there are forces all through the liquid.

Pressure is equal to the ratio of a force and an area taken perpendicularly to this force. But often, in referring to pressures in the interior of a liquid, nothing is said of an area. The area may be imagined, for example, by inserting an imaginary sheet of paper. It is in this sense that it is said that the weight of a liquid produces pressures at all places in the liquid.

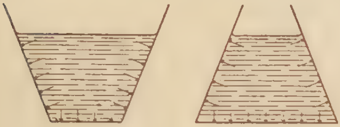


FIG. 27

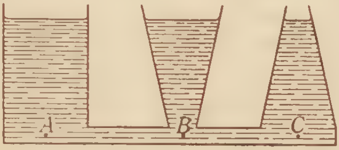


FIG. 28

One of the statements in this section, fact 2, is that pressure increases with depth. This statement is incomplete. The increase of pressure is *directly proportional* to the increase in depth. This is more fully discussed in section 24, where it is proved that the difference in pressure between two levels in water is

$$p = 0.433\,h \text{ pounds per square inch,} \qquad (2)$$

where h is the vertical distance in feet between the two levels. This means that for each foot of increase in depth, pressure increases 0.433 pound per square inch.

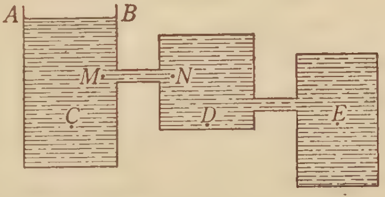


FIG. 29

Fig. 27 shows cases of pressure acting perpendicularly to the surface.

Fact 6 needs some further consideration. It should be obvious, in the case illustrated in Fig. 28, that the pressures at the points A, B, and C must be the same. If they were not, the side pressure at the point of greater pressure would cause a flow of the liquid toward the region of lower pressure. But in the case of some intricately connected system, such as that given in Fig. 29, it is not so evident that the pressures at the points C, D, and E on the same level are equal. The student should try to prove that the pressures are equal in this case (equation (2) may be used).

24. Derivation of the pressure-depth formulas. The principle underlying the theoretical proof of the pressure-depth formula, equation (2), is that pressure in a liquid is produced by the weight of that part of the liquid which lies above the place under consideration. Since pressure is numerically equal to the force acting on 1 square inch, the pressure on the bottom of a vessel is equal to the weight of a column 1 square inch in cross section extending up to the surface of the liquid. For example, the pressure of the water at the bottom of the vessel in Fig. 30 is equal to the weight of a column A of unit cross section; the pressure at the level MN is equal to the weight of a column B of unit cross section. The difference between the pressure at the bottom of the vessel and that at the level MN is equal to the weight of a column C, also of unit cross section. In general, the difference in the downward pressures at two different levels is equal to the weight of a column of liquid of unit cross section and of a height equal to the difference of levels of the two places. Since the weight of a column of water 1 square inch in cross section and 1 foot high is 0.433 pound, the difference in pressure in pounds per square inch in water is given by (2),

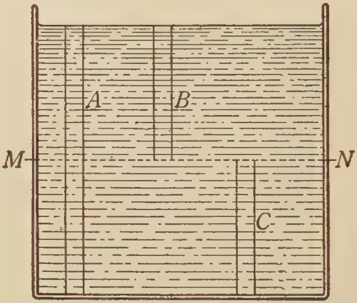


FIG. 30

$$p = 0.433\,h \text{ pounds per square inch,}$$

where h is the difference in the depth measured in feet.

In the C.G.S. system a pressure of 1 dyne per square centimeter is called a *bar*. The weight of a column of water 1 centimeter in cross section and 1 centimeter high is 980 dynes. Hence if p is the increase in pressure in dynes per square centimeter, or bars, and h the difference in depth in centimeters,

$$p = 980\,h \text{ bars.} \qquad (3)$$

If the liquid is one that does not weigh the same as water, its *specific gravity* must be taken into account. The specific gravity of a liquid is defined as the *ratio of the weight of the liquid to the weight of an equal volume of water.* If d is the specific gravity, the weight of a column of the liquid is d times the weight of a column of water of the same height and cross section. Hence for a liquid the specific gravity of which is d, the increase in pressure in pounds per square inch is given by

$$p = 0.433\,hd, \tag{4}$$

and in bars by

$$p = 980\,hd, \tag{5}$$

where h is measured in feet in (4) and in centimeters in (5).

Equations (4) and (5) give the numerical value of the increase in pressure for a change in depth. If the liquid has a free surface (that is, if it is in an open vessel), these equations will give the total pressure due to the weight. In this case h is the depth below the free surface. For example, in Fig. 29 the pressure at D or E is given by these equations if h is regarded as the depth of these points below the free surface AB of the first vessel.

Downward pressure was used in deriving the formulas for computing the numerical value of pressures; but since pressures are the same for all directions, these equations give the numerical value of pressure in any direction.

Since the pressure at any place depends on the depth below the level of the free surface, this distance is often used to express the pressure. When spoken of in terms of depth, the pressure is often called the *head*. For example, if the free surface of a reservoir is 100 feet above the given point, the head of water in the mains at that point is 100 feet (if the water is at rest). Equation (2) gives the numerical relation between head in feet (h) and pressure in pounds per square inch. For example, a head of 92 feet corresponds to a pressure of approximately 40 pounds per square inch. Often the pressure is given in terms of the head when no free surface exists. In that case the head is the height to which the pressure would raise water if an open tube were run up to that height.

25. The hydrostatic paradox. Consider the vessels shown in Fig. 31, which have the same area of base and which are filled with water to the same height. The pressures at the bottoms of these vessels will be equal, for pressure depends only on the depth. Since the areas are equal, the total downward force on the bottom of each is the same. If the vessels when empty have the same weight, we might contend that since the total downward force on

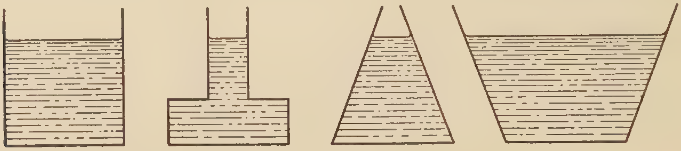


FIG. 31

the bottom of each is the same when filled with liquid to the same height, they should weigh the same when placed on the scalepan of the balance. Where is the fallacy?

26. Units for measuring pressure. In English-speaking countries the *pound per square inch* is commonly used to measure pressure. In some cases *tons per square foot* is used. The unit adopted by the International Congress of Physicists in Paris in 1900 is the *barye,* which is equal to 1 *dyne per square centimeter.* In English the word *barye* is commonly abbreviated to *bar.* Since the bar is a very small pressure, the *megabar,* or 1,000,000 bars, is sometimes used. Pressure is often stated in terms of feet or inches of water, or in inches or centimeters of mercury, the quantity being the height of a column of the liquid the weight of which produces the pressure.

The following gives the numerical relations between the more common units:

1 in. of mercury = 0.491 lb. per square inch
1 cm. of mercury = 0.1934 lb. per square inch
1 cm. of mercury = 0.0133 megabar
1 ft. of water = 0.433 lb. per square inch
1 in. of water = 0.0361 lb. per square inch
1 cm. of water = 0.0142 lb. per square inch
1 cm. of water = 980 bars
1 lb. per square inch = 0.0690 megabar
1 megabar = 14.5 lb. per square inch
1 atmosphere = 1.0132 megabars
1 atmosphere = 14.7 lb. per square inch
1 atmosphere = 1.058 tons per square foot

27. Transmission of pressure; Pascal's principle. Pascal's principle states that *when a pressure is added to a liquid at any place, the pressure everywhere throughout that liquid is increased by an amount exactly equal to the pressure added.* If the pressure in the mains of a city water-supply system be increased at the plant, an increase in pressure is produced over the whole system.

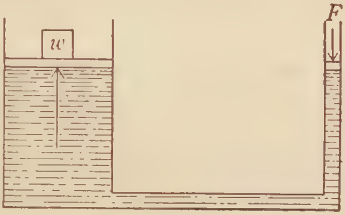


FIG. 32

In the case shown in Fig. 32 the pressure can be increased by a force F applied at the small piston. According to Pascal's principle the pressure throughout the two connected vessels is increased by an amount exactly equal to the pressure applied at the piston. The applied force F will produce a pressure F/a, where a is the area of the piston. This pressure will be transmitted through the water and exert an upward pressure F/a on the large piston. The total force acting upward on this piston is

$$\frac{F}{a}A = F\frac{A}{a},$$

where A is the area of the large piston. If the ratio A/a is large, as it easily may be, large forces can be produced by relatively small ones. The hydraulic press is a machine which uses this method for producing large forces. In practical applications, instead of a simple piston (as shown in Fig. 32) a force pump is often used. The pressure from a pump may be transmitted by pipes to many different places where there are cylinders with large pistons.

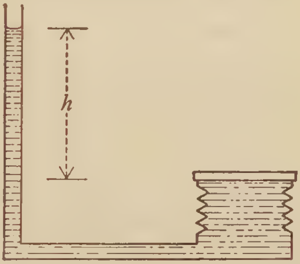


FIG. 33

The *hydrostatic bellows* is another example of the production of large forces by pressure transmitted through liquids. The weight of the water in the side tube (Fig. 33) produces an upward pressure on the underside of the platform of the bellows. If h is the height in feet of the surface of water in the side tube above the underside of the platform, and A the area of the platform, it follows from equation (2) that the total upward force is

$$0.433\ hA \text{ pounds.}$$

It does not make any difference what the diameter of the side tube is, for the pressure depends only on the height of the water.

In Chapter IV it is shown that the atmosphere exerts a pressure on the free surface of liquids. It follows from Pascal's principle that the pressure everywhere throughout a liquid that has a free surface is increased by an amount equal to atmospheric pressure. However, it is customary in referring to the pressure in a liquid or gas to state the amount the pressure is in excess of that of the air. A pressure of 40 pounds per square inch in a water pipe or steam pipe means a pressure of that amount in excess of atmospheric pressure.

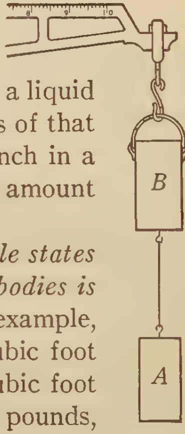


FIG. 34

28. Archimedes' principle. *Archimedes' principle states that the buoyant force on floating or submerged bodies is equal to the weight of the liquid displaced.* For example, if any nonporous material with a volume of 1 cubic foot is completely immersed in water, it displaces a cubic foot of water. Since a cubic foot of water weighs 62.4 pounds, this principle states that the buoyant force is 62.4 pounds. The body apparently loses an amount of weight equal to 62.4 pounds. This principle may be experimentally verified as well as illustrated by the following experiment. The apparatus is illustrated in Fig. 34. A is a brass cylinder which has the same dimensions as the interior of a brass bucket B, for the cylinder when placed inside the bucket will be found to fit it snugly. The cylinder and the bucket are suspended from one side of a balance, and a mass having an equal weight is placed on the other. When the cylinder, which hangs below the bucket, is immersed in water, this side of the balance loses weight. Equilibrium is exactly restored when the bucket is filled level full with water. This shows that the loss of weight is exactly equal to the weight of water in the bucket. Since the volume of the bucket is equal to the volume of the submerged cylinder, the loss of weight is equal to the weight of water having a volume equal to that of the cylinder.

This principle was discovered by the Greek philosopher Archimedes. It is related that Hiero, king of Syracuse, had ordered a crown made of pure gold. Suspecting that the artificers had alloyed it with silver, he asked Archimedes to find out whether the gold was pure or not. It is further related that Archimedes conceived a solution through observing the buoyant force on his body while in a bath. He probably made an intelligent guess, and afterwards, by experiment, verified the principle.

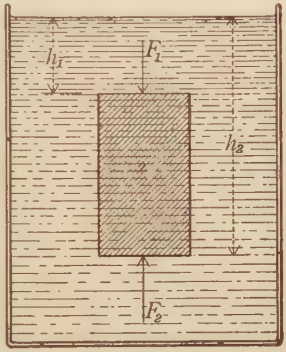


FIG. 35

29. Theoretical proof of Archimedes' principle. Let the submerged body be represented by the shaded area in Fig. 35. On the upper surface of this body there is a downward pressure p_1 and a downward force

$$F_1 = Ap_1,$$

where A is the area of the upper surface of the body. On the lower surface there is an upward pressure p_2 and an upward force

$$F_2 = Ap_2.$$

The resultant upward force is

$$F_2 - F_1 = A(p_2 - p_1).$$

The difference in pressure, $p_2 - p_1$, is the weight of a column of liquid of unit section and of height $h_2 - h_1$. Hence $A(p_2 - p_1)$ is the weight of a column of the liquid of area A and height $h_2 - h_1$; that is, of the same dimensions as the submerged solid. Hence the buoyant force equals the weight of an amount of liquid equal in volume to the volume of the submerged body.

The student will see that this proof is not general,—that it applies only to a body of uniform cross section. However, if he has learned in geometry that an irregularly shaped body may be regarded as made up of a very large number of small, regularly shaped volumes, he should have no serious difficulty in making this proof general.

30. Examples of Archimedes' principle. A ship floating on water has a buoyant force acting on it equal to the weight of the ship. From the principle of Archimedes it follows that the weight of the displaced water is equal to the weight of the ship. Fishes can change the buoyant force acting on them by changing the volume of the displaced water. They do this by means of the "swimming bladder," which can be expanded or compressed at will. A submarine contains compartments which, at the direction of the commander, can be filled with sea water or with air from compressed-air tanks. Thus the amount of the displaced water changes, and the buoyant force decreases or increases as desired. In the case of the "Cartesian diver," a small hollow-glass image of an imp is partly filled with water. Enough air is left inside to make it just float in a jar of water. If one presses on a rubber membrane which is stretched over the mouth of the jar, the pressure is increased throughout the liquid, and the glass imp sinks. Why?

31. Specific gravity; density. *The specific gravity of a body is defined as the ratio of the weight of that body to the weight of an equal volume of water.* Iron is said to have a specific gravity of 7.8, because a piece of iron weighs 7.8 times as much as an equal volume of water.

Sometimes in practical work the number of *pounds per cubic foot* is used instead of specific gravity. If one knows that water weighs 62.4 pounds per cubic foot, and also knows the specific gravity of the substance, the number of pounds per cubic foot can be computed, for it is equal to 62.4 times the specific gravity of the substance. Thus the number of pounds per cubic foot of iron is about 490.

The numerical value of the specific gravity of a substance is independent of the system of units that is used. Why?

Density is defined as follows:

$$\text{Density} = \frac{\text{mass of a body}}{\text{volume of the body}}. \tag{6}$$

From this it follows that density* is equal to the *mass per unit volume of the substance*; for example, in the C.G.S. system it is

---

*If the body is not of uniform density, these definitions of specific gravity and density give the *average* values.

the number of grams in a cubic centimeter. The numerical value of the density of a substance depends on the units used. Since 1 cubic centimeter of water has a mass of almost exactly 1 gram, a volume of water is for all practical purposes numerically equal to its mass when C.G.S. units are used. Hence in these units the density of water is unity; but in another system of units, where such a simple relation between the units of volume and those of mass does not exist, the density of water is equal to a different number.

When measured in C.G.S. units the density of any substance has practically the same numerical value as specific gravity. This can be proved as follows:

Since the volume of a substance in cubic centimeters is numerically equal to the mass in grams of an equal volume of water, the definition of density can be stated for C.G.S. units as

$$\text{Density} = \frac{\text{mass of a body}}{\text{mass of an equal volume of water}}.$$

The relation between the mass and the weight of a body will be fully discussed later. Here it is only necessary to know that the weights of bodies are proportional to their masses. The ratio of the masses of two bodies is the same as the ratio of their weights. Hence when density is measured in C.G.S. units, it is true that

$$\text{Density} = \frac{\text{weight of a body}}{\text{weight of an equal volume of water}}.$$

But this is what we have called the specific gravity; hence, in the C.G.S. system, density is numerically equal to specific gravity. Since the term *density* is used in this text only with C.G.S. units, the density of any particular substance will always be numerically equal to its specific gravity.

32. Measurement of the specific gravity of solids. A simple method of measuring the specific gravity of a solid body is to weigh the body and then immerse it in a vessel *full* of water. The overflow of water is caught and weighed. This overflow has a volume equal to the volume of the body submerged. Hence the ratio of the weight of the body to the weight of the overflow gives the specific gravity.

When the body is of regular dimensions, so that its volume can be determined from measurements, one may proceed as follows: The mass in grams is determined by weighing, and the volume in cubic centimeters is computed from the dimensions. The ratio of these gives the density, and this is numerically equal to the specific gravity.

Irregularly shaped bodies can have their specific gravity determined by the use of Archimedes' principle. The loss of weight of the body when submerged is equal to the weight of an equal volume of water; hence the specific gravity of a body is equal to the weight of the body divided by its loss in weight when submerged. If the student has not already actually performed this experiment in the laboratory, he should think about it until he understands it clearly.

SPECIFIC GRAVITIES

Aluminum	2.70	Lead	11.37
Brass	8.4–8.7	Mercury (15° C.)	13.56
Carbon		Oil (castor)	0.97
diamond	3.52	automobile engine	0.88
graphite	2.3	fine machine	0.86
Coal		Platinum	21.5
anthracite	1.3–1.7	Silver	10.5
bituminous	1.2–1.4	Tin	7.29
Concrete	2.3–2.5	Tungsten	18.8
Copper	8.80–8.95	Turpentine	0.87
Gasoline	0.72–0.75	Wood	
Glass		beech	0.7–0.9
common	2.4–2.6	cypress	0.55
flint	2.9–4.5	elm	0.57
Glycerin	1.26	hickory	0.89
Gold	19.32	mahogany	0.58
Iron		oak	0.7–0.9
pure	7.86	pine (white)	0.38–0.4
gray cast	7.0–7.1	pine (yellow)	0.54–0.72
white cast	7.6–7.7	redwood	0.48
wrought	7.8–7.9	Zinc	7.1
steel	7.7–7.9		

33. The specific gravity of liquids. The specific gravity of a liquid which (like oil or mercury) will not mix with water can be determined to a fair degree of accuracy by a very simple method, as follows: Some of the liquid is poured into a U-tube open at both ends. Into one side is poured water. The heights at which the free surfaces stand above the surface of contact of the two liquids are then measured. In Fig. 36 these are marked h_1 and h_2. The ratio of h_1 to h_2 is the specific gravity of the liquid. That this ratio gives the specific gravity can be proved as follows:

The pressure at all points on a level drawn through the surface of contact of the two liquids is the same in both tubes. Using the formulas for the pressure, as given in equations (2) and (4), the pressure due to the water is

$$0.433\,h_1 \text{ pounds per square inch,}$$

and that on the other side is

$$0.433\,h_2 d \text{ pounds per square inch,}$$

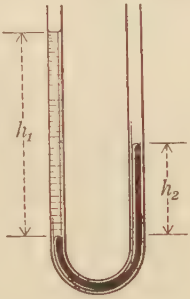


FIG. 36

where d is the specific gravity. Since these two pressures are equal we have

$$0.433\ h_2 d = 0.433\ h_1,$$

or

$$d = \frac{h_1}{h_2}.$$

The two vertical tubes do not need to have the same internal diameter. Why?

If the liquid to be tested mixes with water, the specific gravity can be measured by the apparatus shown in Fig. 37. When some of the air from the upper part of the tubes is withdrawn, the liquids will rise in the tubes. Here, as in the last case, the two columns produce the same pressure, and the specific gravity is equal to the ratio of the heights of the two columns.

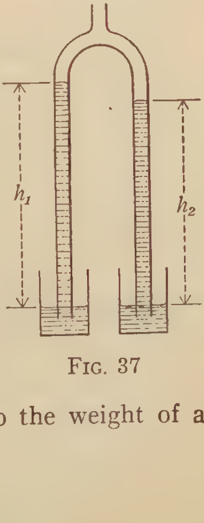


FIG. 37

Very accurate measurements of the specific gravity of liquids have been made by weighing a quartz sphere in different liquids. Since the loss of weight of the sphere is always equal to the weight of a volume of the liquid equal to the volume of the sphere, a method of comparing the weights of equal volumes of liquids is available.

Where only a fair degree of accuracy is needed, but where it is of advantage to get results quickly, *hydrometers* are generally used to determine the specific gravity of liquids. These are weighted glass tubes (Fig. 38) with the upper part graduated. A hydrometer will sink down into the liquid until the buoyant force is equal to the weight of the hydrometer. In heavy liquids it will not sink so deeply as in lighter liquids. By the aid of the graduated scale this difference is measured. Usually the scales are so graduated that the specific gravity can be read off directly.

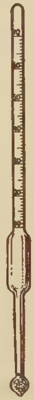


Fig. 38. Hydrometer

34. The compressibility of liquids. In 1762 Canton proved that water was slightly compressible. However, it is only by the application of large pressures that the volume is appreciably reduced. In measuring the compressibility of liquids, experimenters have generally used, as a unit of pressure, the pressure of the atmosphere. Regnault found for water that the compressibility per atmosphere is 0.000048, where the term *compressibility* means the ratio of the change in volume to the original volume. Hence for water the change in volume is equal to

$$0.000048 \times \text{original volume} \times \text{pressure in atmospheres.}$$

The compressibility of sea water is a little less than for pure water. It is equal to 0.000044 per atmosphere. At a depth of 10,000 feet the pressure is about 4400 pounds per square inch, or 300 atmospheres. But 300×0.000044 is equal to 0.0132; hence the specific gravity of sea water at this depth is increased 1.3 per cent. There is no reason for supposing that the density of the sea at great depths is much different from that at the surface.

Mercury is compressed even less than water, its compressibility being about 0.0000038 per atmosphere.

PROBLEMS

1. There is a round hole in the bottom of an otherwise water-tight box. A hollow rubber ball is held on this hole while the box is filled with water. (*a*) Will the ball tend to float? (*b*) Will there be any difference in the result if the ball is quite large or if it is only a little larger than the hole?

2. (*a*) Find the pressure in pounds per square inch due to a column of water 20 ft. high. (*b*) Find the pressure in both bars and pounds per square inch due to a column of mercury 74 cm. high.

3. Compute the pressure in bars and also in centimeters of mercury in each of the following cases: (*a*) the pressure produced by water 200 cm. high; (*b*) the pressure produced by a 50-centimeter column of glycerin, specific gravity 1.26.

4. What must be the minimum pressure inside the city water mains to raise water to an outlet 100 ft. above the mains?

5. The pressure in city gas mains is 2.5 in. of water. Find the pressure in pounds per square inch.

6. The larger piston in a hydraulic press (Fig. 32) is 20 ft. higher than the smaller one. The area of the larger is 100 sq. in., that of the smaller 2 sq. in. If a force of 20 lb. is applied to the smaller one, how large a weight can be lifted by the larger one?

7. The inside dimensions of a jar are 5 cm. in diameter and 30 cm. in depth. It is filled halfway with mercury, and the rest with water. Find (*a*) the pressure halfway down, (*b*) the pressure at the bottom, and (*c*) the total force on the bottom.

8. A diving-bell is 50 ft. below the surface of water. What is the pressure per square inch produced by the water?

9. A rectangular tank measures inside 120 by 240 cm. It is filled with brine (density $= 1.05$ gm./cm.^{3}) to a height of 500 cm. Find the force on the bottom.

10. The specific gravity of aluminum is 2.70. Find the weight per cubic foot.

11. The weight per cubic foot of a certain grade of pressed brick is 150 lb. Find the average specific gravity of the brick.

12. Copper has a specific gravity of 8.9. How many grams of copper must be suspended in water in order that the apparent mass in water shall be 100 gm.?

13. A solid body weighs 100 gm. in air and 64 gm. in a liquid of specific gravity 1.2. What is its specific gravity?

14. A ferry barge with vertical sides has a bottom 40 by 15 ft. When several motor cars were on board, it floated 4 in. deeper. How much did the cars weigh?

15. How much force will it take to sink an oak plank in water if it weighs 25 lb. in air and if its specific gravity is 0.8?

16. An alloy of gold and silver is found to have a specific gravity of 16. Find the percentage of weight of silver in the alloy. The specific gravity of pure gold is 19.32, that of pure silver is 10.5.

17. A rectangular block of cast iron, specific gravity 7.6, is 12 in. long, 8 in. thick, and 9 in. high. (*a*) Find the weight of water it displaces when completely submerged. (*b*) Find the weight of the iron while under water.

18. A body weighs 18.4 oz. in air, 13.9 oz. when submerged in water, and 14.4 oz. when submerged in a second liquid. Compute (*a*) the specific gravity of the substance and (*b*) the specific gravity of the second liquid.

19. How many pounds of wood (specific gravity = 0.6) must be fastened to 5 lb. of copper (specific gravity = 8.9) in order that the combination may barely float in water?

CHAPTER IV

GASES

The weight of air, 35. The pressure produced by the weight of air, 36. Pressure is exerted in all directions, 37. The molecular theory of gases, 38. The barometer, 39. Types of barometers, 40. The pressure of the atmosphere varies with elevation, 41. Pressure gauges, 42. The buoyancy of air, 43. The specific gravity of gases; relative density, 44. The compressibility of gases: Boyle's law, 45. Simple water pumps, 46. Simple air pumps, 47. The siphon, 48.

35. The weight of air. Often one is surprised when he learns for the first time how much weight air really has. A cubic foot of air weighs about 1.3 ounces, a cubic yard about 2.2 pounds. A classroom 30 by 30 by 12 feet high contains about 880 pounds.

It is very easy to show experimentally that air has weight. If a flask or bottle of moderate size is weighed before and after the air has been pumped out of it, the change in weight can be observed on an ordinary balance; or, by the aid of a compression pump (for example, a bicycle pump), air can be compressed into a flask, and the increase in weight of the flask observed.

36. The pressure produced by the weight of air. The earth is surrounded by an immense ocean of air, the total weight of which is so great that the numbers used to express it are too large to have much meaning. The average pressure on the earth's surface produced by the air is the total weight divided by the area of the earth's surface. It is about 14 or 15 pounds per square inch. The total downward force on the floor of a room 15 feet square is several hundred tons. (Must the floor be made stronger on this account?)

The existence of this downward pressure can be shown in a number of different ways. If a rubber membrane is fastened over the open top of a bell jar and the air is partly removed by a filter pump or an ordinary air pump, the membrane is forced downward into the bell jar (Fig. 39). Or, instead of the membrane, the hand can be placed over the open mouth, and the weight of the air felt. Even after feeling it, it is hard to realize that it is the weight of the air above that produces the force. Such experiments as this are often popularly explained as due to the "suction" of the air pump. A little thought will convince one that there can be no force of suction,—that it is the *presence* of air, not its *absence*, which produces the force.

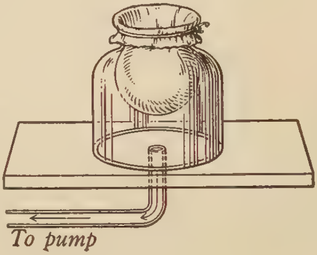


FIG. 39

The actual pressure of the air is seldom constant at any one place for any considerable number of hours. However, the range is only a few per cent—seldom more than 5 per cent. Hence for rough work the pressure of the atmosphere has been used as a unit, or standard, for measuring pressures.

37. Pressure is exerted in all directions. In the experiment of Fig. 39 why is it that the downward pressure is not shown until air is removed from underneath the hand? It is because air in the bell jar exerts an upward pressure on the hand. When part of the air is removed and the upward pressure is decreased, then the downward pressure becomes greater than the upward, and the *difference* is observed. In section 36 a question was left unanswered. Must a floor be made stronger on account of the downward pressure of the air? The answer is No, provided that the air in the room below is at ordinary atmospheric pressure. If the room below should contain a partial vacuum, then the floor would probably need reënforcing.

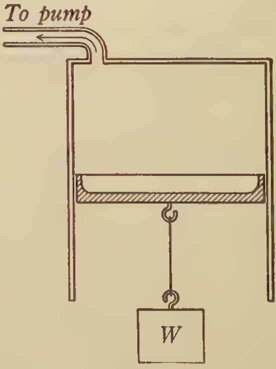


FIG. 40

The upward pressure of air can be shown by a simple, direct experiment. A cylinder, sometimes called a *cymometer*, or a *seven-in-one apparatus*, contains a piston as shown in Fig. 40. Attached to this piston is a heavy weight *W*. When the air inside is partially removed by an air pump, the pressure of the air outside will force the piston up, lifting the heavy weight.

Another familiar experiment is that performed with a tumbler full of water and a card. The card is placed on the tumbler and held there while the tumbler is quickly inverted. The upward pressure of the air will hold the card in place and prevent the water from spilling.

The Magdeburg hemispheres were invented by Otto von Guericke, about 1654, to show the pressure of air. In Fig. 41 is shown a small modern form for classroom use. The halves of the sphere have their joining edges ground true, so that when covered with vaseline or some other lubricant they form a joint which is nearly air-tight. When the hemispheres are placed together and the air is partly removed, they are held together by a considerable force, usually over a hundred pounds. This force is the same, no matter what the position of the hemispheres may be; hence they can be used to demonstrate that air pressure acts in every direction.

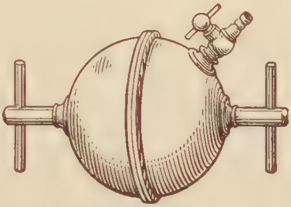


Fig. 41

38. The molecular theory of gases. It is now quite well established that ordinary matter is not a uniform continuous mass but is made up of very minute grains called *molecules*. When the molecules are close together, they exert forces of attraction on each other. In a solid these forces so bind the molecules that they have little freedom of motion, but in a liquid they are not held together in so rigid a way. The forces acting on the molecules are sufficient to keep them from moving apart, so that the volume occupied by the liquid is constant; but the forces are not sufficient to prevent the liquid from changing its shape, or, as is commonly said, from flowing.

In a gas the condition is very different. A gas does not tend to keep a constant volume, but has a tendency to expand. The elastic action of bicycle and automobile tires depends on this fact. Glass bottles tightly sealed may burst when placed in a vacuum. The pressure exerted by steam in boilers or by the ignited gases in gas engines are other examples of this tendency of gases to expand. But why does a gas tend to expand? This is easily explained if we assume that the molecules are relatively far apart and moving continually with high speeds. Under these circumstances the restraining influence of their attraction is too small to overcome the effect of their high speeds, so that they tend to move apart, to fly off in every direction. The molecules in a solid are also supposed to be in rapid motion, but they are so close together that the motion is probably an oscillatory one rather than one which takes the molecule from one part of the body to another, as in the case of a gas. Some idea of the relatively greater distances apart of the molecules of a gas, as compared with those of a solid or liquid, is obtained when we note the difference in the compressibility of gases as compared with liquids or solids. Gases under great pressure are reduced to comparatively small volumes. Under a pressure of 3000 atmospheres the volume of air is reduced to about 1/700 of its original value. Steam when changed back into water occupies only about 1/1600 of its volume. It is believed that the change in volume on compression is due not to changes in the size of the molecules but to changes in the spaces between them.

The kinetic theory of gases gives a simple picture of the way in which a gas exerts pressure when confined in some vessel. Imagine a board struck by a large number of balls, one after another. Practically there would be a constant force acting on the board. A continuous stream of sand blown against a surface exerts just such a pressure. In a gas there are enormous numbers of molecules striking against the walls of the containing vessel, each molecule giving a small impulse. While each of the molecules has a very small mass, yet there are so many of them, and they move with such high speeds, that forces of considerable magnitude are produced. The theory of this phenomenon has been developed to such an extent that it is now possible, if one knows the density of the gas and the pressure exerted by it, to compute the probable average speed of the molecules. This speed at ordinary temperatures is always high; for example, air molecules have an average speed of over 1200 feet per second.

When we understand that the molecules of a gas have such enormous speeds, we have no difficulty in comprehending many facts about gases; for example, why they expand so quickly and why a small quantity of illuminating gas can diffuse so rapidly throughout an entire room.

39. The barometer. The barometer is an instrument for measuring the pressure of the atmosphere. A simple form can be made by using a piece of clean, dry, glass tubing, about three feet long, with one end closed. This tube is filled with clean mercury, care being taken to remove all air bubbles that cling to the walls of the tube. A finger is now placed over the open end, and the tube inverted in a dish of mercury in such a way that air does not bubble up into the tube. The mercury in the tube will fall until its surface stands 29 or 30 inches above the level of the surface of the mercury in the dish (Fig. 42). The height at which the mercury stands in the tube is a measure of the air pressure on the open surface of the mercury.

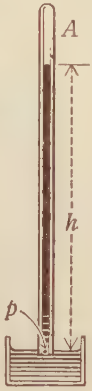


FIG. 42

In many kinds of experimental work it is necessary to know the pressure of the atmosphere; hence the barometer is a common laboratory instrument. The simultaneous observations of air pressure, taken daily over much of this continent, are of great value in predicting weather changes.

The theory of the barometer can be understood by the aid of Fig. 42. Let us assume that there is no air or vapor in the end *A* of the tube. In section 23 it was stated that the pressure in a liquid at all points on the same level must be the same. Hence at a point *p* inside the tube, but on the same level as the free surface of the mercury outside, the pressure must be the same as at points on the free surface. But at the surface of the mercury in the dish the pressure is equal to that of the atmosphere. Hence at *p* the pressure must be that of the atmosphere. It follows that the mercury must extend up to a height *h* sufficient to produce a downward pressure equal to that of the atmosphere. The size of the tube, provided it is not so small that capillary action enters (sect. 164), does not affect this height.

A discussion of the following experiment will aid in understanding the principle of the barometer:

A glass cylinder is partly filled with water and is inverted in a dish of water (Fig. 43). The upper part of the cylinder will now contain air at a pressure p. At the free surface of the liquid in the dish the downward pressure is that due to the air, say P. The downward pressure at a point inside the cylinder on the same level as the free surface of the liquid outside is due to two things: the weight of a column of water of unit section extending up to the top of the water, and the pressure p of the air in the upper end of the cylinder. If the pressure due to the water is p_1, the downward pressure is $p_1 + p$. But this pressure must be the same as that of the free surface outside at the same level; that is, it must be equal to P. The pressure p_1, due to the water, can be computed by the rules given in section 24. If there is no air in the cylinder, and p is hence equal to zero, the water will rise until it is about 33 feet high, provided the cylinder is sufficiently long. The greater height to which water, as compared with mercury, rises is due to the fact that the weight of mercury is 13.6 times that of an equal volume of water.

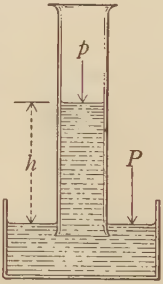


FIG. 43

40. Types of barometers. There are two types of barometers: the mercury and the aneroid. If an accurate instrument is desired, great care must be used in making a mercury barometer. It is usually a matter of considerable patience to get a tube which is closed at one end thoroughly dried out after washing. It is also difficult to get all the air out of the tube. If there is enough air left in the top of the tube to cause a considerable error, a sharp metallic click will not be obtained when the tube is tipped so far over that the mercury strikes the upper end. (When the tube is tipped, the upper surface of the mercury should stay at the same level. Why?)

The *aneroid* barometer contains a hollow, hermetically sealed metal box in which there is a partial vacuum (Fig. 44). When the pressure of the atmosphere increases, the top of this box is bent in; when the pressure decreases, it springs out. By a system of levers a needle is made to move back and forth over the face of the dial as the top of the box moves in and out. The dials are usually graduated to read inches of mercury.

The height of the column of mercury depends not only on the pressure of the atmosphere but also on temperature (mercury expands with an increase in temperature; hence a column of a certain height does not weigh as much as when colder). It is customary to reduce barometer readings to 0° C., or 32° F.; that is, to state what the height would have been if the mercury had been at a temperature of 0° C., or 32° F.

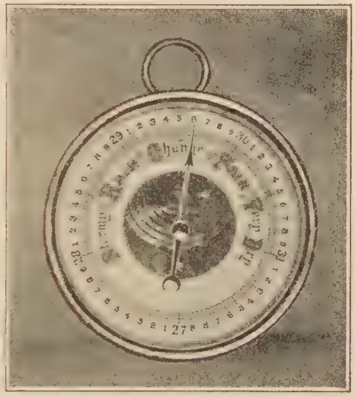


FIG. 44. An aneroid barometer

41. The pressure of the atmosphere varies with elevation. The pressure of the atmosphere, as we have learned, is due to the weight of the air above us. At a high elevation there is less air above, and hence a smaller pressure. The air pressure at an elevation of 20,000 feet is only half of that at sea level, for about half of the mass of the entire atmosphere is below that elevation. At an elevation of 900 feet above sea level the barometer is about 1 inch lower than at sea level. (The student should make an effort to find out his elevation and the approximate local correction of a barometer.)

Portable aneroids are often made with scales graduated for measuring elevations. If readings are made both in a valley and on a hill, the height of the hill above the valley can be determined. Aviators use barometers to determine their heights. Commercial types of aneroid barometers are made which are so sensitive that they will indicate changes in atmospheric pressure between the floor level and the top of an ordinary table.

In weather forecasting it is customary to correct all barometer readings for elevation as well as for temperature. The reading is changed to what it would have been if the station had been at sea level and if the temperature of the mercury had been 0° C., or 32° F. The barometric pressure shown on the weather maps issued by the United States Weather Bureau represents always the corrected values.

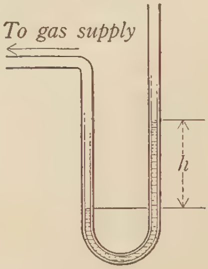


FIG. 45

42. Pressure gauges. The principle of one type of pressure gauge, the *open-tube gauge*, can be shown by the following experiment: Some water is placed in a bent glass tube (Fig. 45), and one end of the tube is connected to the city gas pipes by rubber tubing, the other end being left open to the air. When the gas is turned on, the water falls on one side and rises on the other. The *difference* between the pressure of the gas in the mains and that of the air is equal to the weight of a column of water of unit cross section and of height equal to *h*, the difference in level of the two sides. It is customary to express the pressure in gas mains in inches of water. To reduce to other units use the values given in section 26.

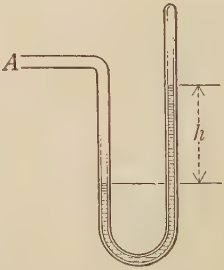


FIG. 46

The *closed-tube gauge* (Fig. 46) contains air under pressure, and the bent part contains some liquid—mercury, oil, or water. The pressure to be measured compresses the inclosed air to a degree depending on the pressure. In this method the total pressure is measured, and not the difference between it and atmospheric pressure.

One form of *vacuum gauge* is constructed in the same way except that (1) water or any other liquid which readily evaporates cannot be used and (2) the closed end contains no air. When the end *A* (Fig. 46) is open to the air, the closed end is entirely full of the liquid, which is held there by atmospheric pressure. If *A* is connected to a vessel, and the pressure in it reduced low enough, the liquid in the closed end will fall. If *h* is the difference in level of the liquid in the two arms, the pressure in the vessel is equal to the pressure produced by a column of the liquid of height *h*; or, more conveniently stated, the pressure in the vessel attached to A is equal to h inches, or h centimeters, of the liquid.

In the *Bourdon spring gauge* the liquid or gas the pressure of which is to be measured is connected to the interior of a hollow bent metal tube of elliptical cross section. When the pressure inside this tube is raised, the tube tends to straighten out and moves an index over a scale. Gauges of this type are commonly used for measuring pressures greater than 1 atmosphere, such as the pressure of steam and of water.

43. The buoyancy of air. Archimedes' principle, which states that a body immersed in a fluid is buoyed up with a force equal to the weight of the fluid displaced, applies to gases as well as to liquids. A hot-air balloon ascends because the balloon and contents weigh less than the colder air which it displaces. Large airships are filled with as large a percentage of hydrogen as possible, for hydrogen is the lightest of all known gases. Helium may be used, for it also is much lighter than air. Ordinary bodies are buoyed up by the air with small but appreciable forces. Since a cubic foot of air weighs 1.3 ounces, the buoyant force in ounces equals the volume of the body in cubic feet times 1.3. A man of average size is buoyed up with a force of about 3 ounces.

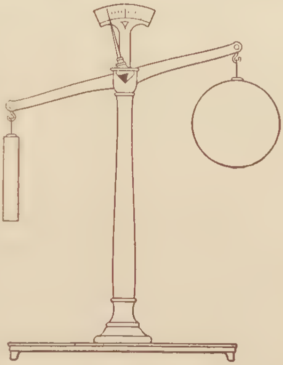


FIG. 47

In accurate weighing, corrections must be made for this buoyant force. If the weights used are of brass, no error is made in weighing brass objects; but in weighing water the error is about one tenth of 1 per cent, or 1 part in 1000.

A lecture experiment often shown to illustrate this buoyant force is as follows: A small balance (Fig. 47) carries a solid brass weight on one side and a hollow brass ball on the other. When this balance is placed under the bell jar of an air pump and the air removed, the ball, which appeared lighter in the air, will now seem heavier than the other weight. (This experiment will work only when the hollow ball is hermetically sealed. The student should attempt to explain this.)

It is said that Aristotle weighed a skin twice: when it was full of air and when it was empty and collapsed. Since the two weights appeared the same, he concluded that air had no weight. What is the fallacy?

In section 35 it was stated that the fact that air has weight could be established by the experiment of weighing a flask when empty and then when full of air. But this experiment is performed with the flask, balance, and weights immersed in the air. Imagine a similar experiment performed with all the apparatus under water. An empty bottle is weighed, and the bottle is again weighed after the stopper has been pulled out and water admitted. The weight of the bottle is now apparently greater. Why? Because there is a decrease in the amount of water displaced, and hence a decrease in the buoyant force acting on the bottle. The same bottle broken to fragments will weigh the same on the completely submerged balance as the bottle of water. After the student has carefully thought out the details of the operation of the submerged balance, he will be able to understand that the experiment of weighing a flask when empty and then when full of air shows only indirectly that air has weight. The experiment, also referred to in section 35, of weighing a flask into which air has been compressed shows the weight of air more directly.

44. The specific gravity of gases; relative density. The specific gravity of liquids and solids is defined with water as a basis. But in the case of gases air is usually taken as the standard, so that specific gravity is the ratio of the weight of gas to the weight of an equal volume of air, care being taken to keep the pressures and temperatures at some standard value. Usually the standard pressure is 76 centimeters of mercury, and the standard temperature the melting point of ice, which is 0° C., or 32° F.

Density is defined in section 31 as the ratio of mass to volume. The *relative density* of a gas is defined as the ratio of the density of the gas to the density of air. The relative density is numerically the same as the specific gravity referred to air as a basis.

SPECIFIC GRAVITY OR RELATIVE DENSITY OF GASES AT 0° C. AND 76 CENTIMETERS OF MERCURY

Air	1.000	Hydrogen	0.0695
Argon	1.377	Nitrogen	0.968
Carbon dioxide	1.529	Oxygen	1.105
Helium	0.1381	Saturated steam at 100° C.	0.462

DENSITY OF GASES IN GRAMS PER CUBIC CENTIMETER AT 0° C. AND 76 CENTIMETERS OF MERCURY

Air	0.001293	Hydrogen	0.000090
Ammonia	0.000771	Nitrogen	0.001251
Carbon dioxide	0.001977	Oxygen	0.001429

45. The compressibility of gases: Boyle's law. Increasing the weight on a piston, such as is illustrated in Fig. 48, will cause the volume of the gas in the cylinder to decrease until the increased pressure of the gas is sufficient to support the piston and the weight. The question is, How much does the volume of a gas change with a given change in pressure? This was answered in England by Robert Boyle, about 1662, and independently in France by Edme Mariotte, about 1676. They showed that when the temperature is kept constant, *the volume of a gas varies inversely as the pressure.* Increasing the pressure of a gas threefold, the volume is reduced to one third; or, if the pressure is made half as great, the volume becomes twice as large. As a numerical illustration of this law consider the following case: An automobile tire is pumped up to a pressure of 75 pounds per square inch (greater than atmospheric pressure). What is the decrease in the volume of the air? If we call 75 pounds per square inch 5 atmospheres (its approximate value) the total pressure in the tire is 6 atmospheres. By the law discovered by Boyle the volume of the air will be one sixth of what it was before compression, or the density in the tire is six times as great as that of the atmosphere.

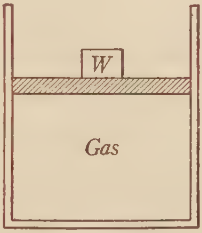


FIG. 48

Boyle's law can be stated as follows: *At a constant temperature the product of pressure and volume of any gas is constant.* Or, when the temperature is constant,

$$p_1v_1 = p_2v_2,$$

where p_1 and v_1 are the pressure and volume at one time, and p_2 and v_2 those at another time.

Boyle's law is only an approximation, but it is close enough for most practical purposes. For very high pressures the error in the law becomes large.

46. Simple water pumps. All pumps which stand at a higher elevation than the level of the water supply use the pressure of the atmosphere. When the piston of the ordinary lift pump (Fig. 49) is raised, the air below the piston expands and its pressure decreases. Then the larger pressure in the pipe below opens the inlet valve A, and some air flows into the cylinder. Atmospheric pressure on the water in the well then raises the water part way up in the pipe. At the end of the upstroke the weight of the inlet valve closes it. On the downstroke the air below the piston is compressed; and when its pressure becomes greater than atmospheric pressure, the outlet valve, B, in the piston is opened and some of the air flows through the piston. The outlet valve will close itself when the piston reaches the lowest point. The operation is then repeated, the water rising higher on each stroke until it flows out of the pump. After the pump becomes full of water, the operation is easily understood. But it must be clearly borne in mind that it is the pressure of the atmosphere which raises the water in the pipe, the piston removing the effects of the downward pressure of the air on the inside of the pipe.

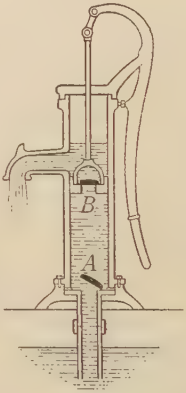


FIG. 49. The lift pump

There is a limit to the height a pump should be placed above the water level. If the construction of the pump were sufficiently accurate so that a good vacuum could be produced in the cylinder, water could be raised about 33 feet. In ordinary pumps the piston must be much nearer to the water level than that. Often the valves and the piston are so loosely fitted that they must be "primed"; that is, covered with water to make them work.

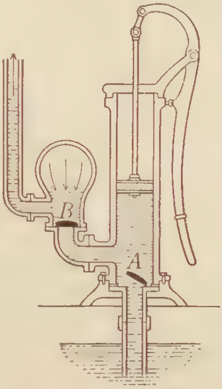


Fig. 50. The force pump

The lift pump is convenient for many purposes, but it is not suitable for giving pressure to the water. One type of *force* pump is illustrated in Fig. 50. It contains an inlet valve, *A*, at the bottom of the cylinder, and an outlet valve, *B*, in the side-connected pipe. In this type the water, which rises into the cylinder on the upstroke, is forced out through the side pipe by the downward stroke of the piston. When the piston is moving downward, part of the water flows into the air chamber, which is above the valve *B*, thus compressing the air. This air, expanding during the upstroke of the piston, drives the water out of the cylinder, thus tending to make the flow from the pump more uniform.

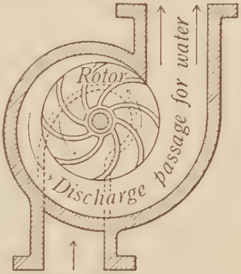


Fig. 51

Double-acting force pumps are often used in pumping-stations. A double set of valves is placed so that the water is pumped from both sides of the piston.

Centrifugal pumps are useful where large volumes are to be moved under low pressures. The water enters at the center of the revolving blades (Fig. 51). Commercial air blowers are frequently a form of centrifugal pump. Another type is similar to an electric fan, the revolving blades cutting through the air with a screw motion. The screw propellers of an airplane or a boat are of this type. The air or water is driven in one direction, and the airplane or boat in the opposite direction.

47. Simple air pumps. The common vacuum air pump is the same type as the lift pump described in the last section. The position of the valves *o* and *i* is shown in Fig. 52. When the piston moves upward, the pressure in the cylinder is decreased so that the air expanding under the bell jar opens the inlet valve, some of the air flowing into the cylinder; when the piston moves downward, the air in the cylinder is compressed until it can open the valve in the piston and flow out. The next stroke takes out more of the air in *r*, and so on. A "vacuum" with a pressure lower than 2 or 3 millimeters of mercury cannot be obtained with this type. The reason for a limit is easily understood. Although the valves are made very light, some pressure is required to open them. In an improved form the valves are opened and closed mechanically by some attachment connected with the piston or another moving part. Another serious difficulty is that when the piston is at the bottom of the cylinder, there will be some air left in the crevices under or around the piston. This air will be at atmospheric pressure, but it expands when the piston is raised. No pressure lower than the pressure of this expanded air can be obtained with the pump. In some forms all the excess space is filled with oil, and much better vacua are obtained. Sometimes two pumps are connected in series and operated at the same time. The first pump produces a partial vacuum above the piston of the second. This greatly increases the effectiveness.

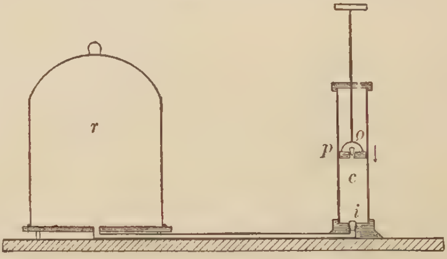


Fig. 52. Exhaust pump

In recent years there have been developed a number of different types of high-speed pumps which will give very high vacua, but their action is not simple and will not be described here.

Compression air pumps work on the same principle as the water force pumps. For obtaining very high pressures it is customary to use several force pumps in series. The air passes from one pump to another, each increasing the pressure.

48. The siphon. The siphon is used a great deal in carrying liquid from a higher to a lower level when there is some high place between them. In Fig. 53 the siphon is the inverted U-shaped tube open at both ends and entirely full of water. In such a tube the water will continue to flow until it reaches the same level in both vessels. A complete explanation of the theory of the siphon involves principles of *flowing* liquids which cannot be gone into here. The simplest explanation is one which shows that the liquid in the tube cannot be at rest. If the liquid in the siphon were at rest, then, according to what was learned in section 23, points on the same level in the two arms must have the same pressure. It will be seen that they do not. At the level of the free surface in the lower vessel the pressure both inside and outside of the tube is atmospheric pressure, provided that the liquid is at rest; but at this level in the other vessel the pressure must be much greater than atmospheric. Such inequalities can be easily shown to exist at other levels. It follows that the liquid in this tube cannot be in equilibrium and at rest. It must be in motion.

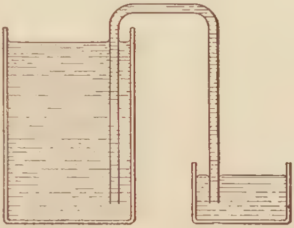


FIG. 53

Water will flow through a siphon which is not more than about 30 feet higher than the level of the source; the maximum height for mercury is about 29 inches. Why?

By means of the *inverted siphon* water in a pipe line may be taken under a river and up on the other side. Water will flow through this pipe without the aid of pumps if the outlet is lower than the place where the water enters the pipe.

PROBLEMS

1. Compute the pressure in bars and also in centimeters of mercury in the following cases: (*a*) 600 cm. under water; (*b*) at the bottom of an open vessel which has mercury 10 cm. deep, with 20 cm. of water on top of the mercury. (The atmospheric pressure is 74 cm. in both cases.)

2. A tube closed at one end is inverted in a dish of mercury, with the mercury standing 20 cm. high in the tube. Find the pressure in the upper part of the tube. The barometer stands 74 cm. high. Reduce the answer to dynes per square centimeter.

3. What is the total weight in pounds of the air in a column 1 sq. ft. in cross section extending from the ground up, when the barometer is 30 in. high?

4. Compute the pressure in bars inside the inverted cylinder of Fig. 43 if *h* is 20 cm. and if the barometer is 74 cm. high. (The liquid is water.)

5. What is the approximate weight of air per cubic foot when the pressure is 75 lb./sq. in. greater than the normal pressure of the atmosphere?

6. If the air in an automobile tire is at a pressure of 80 lb./sq. in. more than atmospheric pressure, what is the ratio of the density to that of air at normal pressure? (Assume that the atmospheric pressure is 15 lb./sq. in.)

7. The total buoyant force on one of the United States army balloons is 6300 lb. when in air at normal pressure and density. What is the volume of the balloon in cubic feet?

8. Show that the lifting-power of a balloon filled with hydrogen is about 7 per cent greater than that of one filled with helium.

9. The area of the piston in an apparatus similar to that shown in Fig. 40 is 20 sq. in. The opening in the upper part is sealed, and a weight of 50 lb. is hung from the piston. Neglecting the effects of friction, find the pressure of the air on the inside of the cylinder. (Assume the atmospheric pressure to be 15 lb./sq. in.)

10. It is found that the pressure in the gas mains on the top floor of an office building is greater than in the basement. Explain.

11. If air at normal pressure has a mass of 1.3 kg. per cubic meter, and hydrogen 0.09 kg., how large must the gas bag of a balloon be if the total load has a mass of 300 kg. and if it is to be filled with hydrogen at normal pressure?

12. From what depth in water has an air bubble risen which doubles its volume when it reaches the surface? (The barometric pressure is 75 cm.)

13. (*a*) What is the apparent loss of mass on account of the buoyant effect of the atmosphere on 1000 gm. of brass? (*b*) on 1000 gm. of mercury? (*c*) If brass weights are used to weigh 1000 gm. of mercury, what will be the magnitude of the error produced by the buoyancy of the air? (The air is at normal pressure and density. Density of brass = 8.4, of mercury = 13.6.)

14. To what height can a liquid of 1.8 specific gravity be raised inside a siphon if the barometric pressure is 76 cm. of mercury?

15. What would be the maximum height over which glycerin, of specific gravity 1.26, can be siphoned when the barometer is 75 cm.?

16. The pressure of air in a vessel as measured by an open-tube gauge (Fig. 45) is 30 cm. of water. Find the ratio of the density of this air to that of the atmosphere if the barometer is 74 cm. high.

17. The closed arm of a gauge such as is illustrated in Fig. 46 contains a vacuum; the liquid is mercury. If h is 2 cm., what fraction of air remains in the vessel to which the gauge is connected, the original pressure being 74 cm. of mercury?

18. When the vessel attached to the gauge of Fig. 46 is at atmospheric pressure, 75 cm., h is 5 cm. Find the pressure in the vessel when h is 30 cm. and the volume in the closed end is half of what it was in the first case. (The liquid in the gauge is mercury.)

CHAPTER V

UNIFORMLY ACCELERATED MOTION

Velocity, 49. Acceleration, 50. Velocity acquired under constant acceleration, 51. Distance traveled in uniformly accelerated motion, 52. Special cases, 53. Geometrical interpretation, 54. Falling bodies, 55. Speed and velocity, 56. Projectiles, 57. Independence of motions, 58.

In some of the chapters which follow there will be constant reference to the motion produced by forces and to the effects of such motions. It will be of great help if some of the simpler concepts and laws of motion are kept clearly in mind. For that reason there is given in this chapter an introduction to the subject of accelerated motion. Special attention should be given by the student to the technical term *acceleration*, for it must be clearly understood in order to avoid confusion later.

49. Velocity. Everyone is familiar with the concept of velocity and with the computations of distance traveled when the time and average velocity are known. The law used for such computations is

$$\text{Distance} = \text{velocity} \times \text{time},$$

or

$$s = vt. \qquad (7)$$

This law applies to two cases: (1) when the velocity v is constant during the time t; (2) when the velocity is changing and v is the average velocity during the time t.

Equation (7) really states a definition of velocity and may be called the *defining equation* of velocity. Such an equation cannot be logically verified by experiment, for it is true because v is always given the numerical value which will satisfy the equation.

50. Acceleration. We are frequently more interested in the *changes* in velocity than in the velocity itself. For example, some types of locomotives are designed to acquire a high speed more quickly than other types. To the designer of such a locomotive the point of interest is how quickly the velocity can be changed, or, in other words, what the *rate of change of the velocity* can be. This *rate* is such an important quantity that a special name has been given to it. It is called the *acceleration*. When the unit of time used is the *second*, this term is defined as follows: *Acceleration is the change per second of the velocity.*

The meaning of this can be clearly understood by studying the data given in the following table. The first column gives the number of seconds after a smooth, heavy ball has begun falling; the second column gives the approximate velocities, in feet per second, that the ball has acquired at the end of the periods given in the first column; the third column gives approximately the total distances passed over by the ball during the corresponding number of seconds.

Time in Seconds	Velocity in Feet per Second	Distance in Feet
0	0	0
1	32	16
2	64	64
3	96	144

The table shows that the velocity is increasing at a uniform rate, for during each second the ball gains a velocity of 32 feet per second. Hence the acceleration is constant and equal to 32 feet per second per second.*

The rate of change of velocity, and the rate of change in distance s, while related, are not the same thing and should not be confused. If the student will compute the rate of change in the distances in the last column of the table, he will find that that rate is not a constant but increases with time.

51. Velocity acquired under constant acceleration. Experiments with falling bodies show that they tend to move with a constant downward acceleration of approximately 32 feet per second per second. What would be the velocity acquired in 3 seconds by a ball thrown downward with an initial velocity of 40 feet per second? Since the gain each second in velocity is 32 feet per second, the gain in 3 seconds is 3 × 32, or 96 feet per second. Obviously

---

* The student should see that time enters twice in computing the acceleration, which is a rate of change of a rate. Since it is customary to state after the numerical value of a quantity the units used, the word *second* appears twice in the above. The student should be able to find a case where the acceleration might be stated as 10 feet per minute per second. Often the units are abbreviated; as, 32 ft./sec.^{2}

$$\text{Final velocity} = \text{initial velocity} + \text{gain in velocity}. \tag{8}$$

Hence $$\text{Final velocity} = 40 + 96 = 136 \text{ ft./sec.}$$

The student should study this example until he understands it thoroughly. He will then understand readily the following statement: If a represents the acceleration, t the time, u the initial velocity, and v the final velocity, the gain in velocity is at, and the final velocity is

$$v = u + at. \tag{9}$$

This is the first of the general laws of uniformly accelerated motion. In equations such as this it is important that the student should be certain that he understands it, that he knows *why* it is true. He should be able to see that equations (8) and (9) mean the same thing, the second one merely stating the first in a more abbreviated way.

52. Distance traveled in uniformly accelerated motion. In the problem of a body thrown vertically downward with an initial velocity of 40 feet per second, how should one proceed to calculate the distance traveled in 3 seconds?

This can be solved in a rather simple way: the initial velocity is 40 feet per second, and the final velocity is $40 + (32 \times 3) = 136$ feet per second. The *average* velocity will be half the sum of the initial and final velocities,* or

$$\text{Average velocity} = \tfrac{1}{2}(40 + 136) = 88 \text{ ft./sec.}$$

But the distance traveled is equal to the average velocity times the time, or,

$$\text{Distance} = 88 \times 3 = 264 \text{ ft.},$$

which is the desired result.

---

* The average velocity is equal to half of the sum of the initial and final velocities only for the special cases where the velocity is changing at a uniform rate.

By following this method a *general* law for computing the distance can be derived. Let the initial velocity be represented by u, the final velocity by v, and the average velocity by $\bar{v}$. The average velocity, as before, will be half the sum of the initial and final velocities, or

$$\bar{v} = \tfrac{1}{2}(u + v).$$

Substituting in this the value of v from equation (9),

$$\bar{v} = \tfrac{1}{2}(u + u + at) = u + \tfrac{1}{2}at.$$

The distance s is given by equation (7),

$$s = \bar{v}t.$$

Hence $$s = (u + \tfrac{1}{2}at)t,$$

or $$s = ut + \tfrac{1}{2}at^2, \tag{10}$$

which is the general law for the distance s traveled in a time t. The student should see that the method of this algebraic process is exactly the same as the numerical one preceding it.

Equations (9) and (10) are the general laws of uniformly accelerated motion and enable one to solve any problem in which not more than two of the quantities a, v, u, t, and s are unknown.

53. Special cases: *Initial velocity zero.* In case the initial velocity is zero (that is, when the body starts from rest), equation (9) for the final velocity becomes

$$v = at, \tag{11}$$

and the distance (equation (10)),

$$s = \tfrac{1}{2}at^2. \tag{12}$$

Decreasing velocity. In case the body is losing velocity at a constant rate a, in t seconds the loss in velocity will be equal to at. Hence the resultant velocity at the end of t seconds is

$$v = u - at. \tag{13}$$

This is equation (9) with the sign of a changed. It is an instance of a general fact that the rate of *decrease* of a velocity is a negative acceleration. Whenever the velocity is decreasing, equations (9) and (10) can be used, but a must have its sign changed. For example, for a negative acceleration, equation (10) becomes

$$s = ut - \tfrac{1}{2}at^2. \tag{14}$$

An example of this sort of motion is the case of a ball thrown vertically upward with an initial velocity of 96 feet per second. Using equations (13) and (14) we can compute the distance it has risen and the velocity at the end of any time t. At the end of 1 second,

$$t = 1.$$

Then, by equations (14) and (13), since $u = 96$ and $a = 32$,

$$s = 96 - \tfrac{1}{2} \times 32 = 80 \text{ ft.},$$
$$v = 96 - 32 = 64 \text{ ft./sec.}$$

At the end of 2 seconds

$$t = 2,$$
$$s = (96 \times 2) - \tfrac{1}{2}(32 \times 2^2) = 128 \text{ ft.},$$
$$v = 96 - (32 \times 2) = 32 \text{ ft./sec.}$$

At the end of 3 seconds

$$t = 3,$$
$$s = (96 \times 3) - \tfrac{1}{2}(32 \times 3^2) = 144 \text{ ft.},$$
$$v = 96 - (32 \times 3) = 0.$$

These results give the height and the upward velocity at different times. With this particular initial velocity it is seen that the ball reaches its maximum height of 144 feet at the end of the third second, for at that instant the velocity is zero. After that the ball falls downward, as shown by the conditions at the end of 4 seconds:

$$s = (96 \times 4) - \tfrac{1}{2}(32 \times 4^2) = 128 \text{ ft.},$$
$$v = 96 - (32 \times 4) = -32 \text{ ft./sec.}$$

The distance 128 feet is not the total distance traveled but the distance of the ball from the starting-point. Note the *negative* velocity. A negative velocity is one in a direction opposite to the positive velocity.

It should be kept in mind that in such computations the effect of the friction of the atmosphere is entirely neglected.

54. Geometrical interpretation. The type of motion that we call uniformly accelerated motion can be given a simple and instructive geometrical interpretation.

In Fig. 54 the distances OP_1, OP_2, and OP_3 represent time intervals equal to 1, 2, and 3 seconds respectively. The distance AO represents the initial velocity of the body. The line P_1N_1 represents the velocity at the end of the first second, M_1N_1 being the gain in velocity during that second. P_2N_2 is the velocity at the end of the second second, and P_3N_3 that at the end of the third second. The steps or increments, M_1N_1, M_2N_2, and M_3N_3, are each numerically equal to the constant acceleration a of the moving body. The distance of any point on the line AB above the line OD represents the value of the velocity v for the corresponding time t. It is left as an exercise for the student to derive equation (9) from this figure.

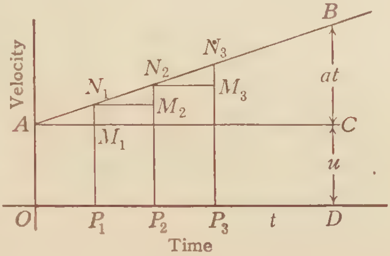


FIG. 54

In the same figure we can find a graphical interpretation of the total distance traveled in a time t. If BD represents the velocity at a time t, and AO the initial velocity, the average velocity $\bar{v}$ is

$$\tfrac{1}{2}(AO + BD).$$

Since the total distance s is the average velocity multiplied by the time,

$$s = \tfrac{1}{2}(AO + BD)OD.$$

The right-hand side of this equation is equal to the area of $OABD$; hence the distance s is numerically equal to the area below the line AB in the figure. Since $AO = u$, and $BD = u + at$, the expression for this area reduces to the form of equation (10),

$$s = ut + \tfrac{1}{2}at^2.$$

55. Falling bodies. The precise laws of falling bodies are very complicated on account of the resistance offered by the air. In a vacuum bodies of all shapes and weights fall with an equal acceleration. This can be experimentally shown by the classic "guinea and feather" tube. A glass tube 4 or 5 feet long contains some light objects, such as pieces of feather or scraps of tissue paper, and also some heavier ones, such as small pieces of metal. The tube has at one end a stopcock which can be connected with an air pump. When a fairly good vacuum is obtained in the tube, and the tube quickly inverted, the various bodies in it are seen to fall with the same speed; but when the air is admitted, and the tube again inverted, the bodies will be seen to fall with very different speeds. Air resistance depends on the shape and size of the body. Moreover, it increases rapidly with the speed of the body. If the falling body is smooth and heavy, air resistance may be neglected for low speeds. In that case the acceleration of a falling body is a constant.

The acceleration of a body falling freely without friction is usually denoted by the letter g. The equations for the velocity and distance of falling bodies are obtained by substituting g for a in equations (9) and (10), or

$$v = u + gt, \tag{15}$$

$$s = ut + \tfrac{1}{2} gt^2. \tag{16}$$

The quantity g varies slightly from place to place over the earth's surface. It depends on both elevation and latitude (see section 142). But for most places the approximate value of 32 feet per second per second, or 980 centimeters per second per second, may be used.

In solving problems of falling bodies special attention should be paid to the signs of the quantities v, u, and g, as explained in section 53.

56. Speed and velocity. A rather technical but simple distinction is sometimes made between the terms *speed* and *velocity*. The term *speed* expresses the numerical value of the rate of motion, but does not contain the idea of direction; on the other hand, the term *velocity* contains the notion of *direction* as well as the *magnitude* of the motion. Any statement of the velocity of a body must be made up of two parts: the speed and the direction.

A velocity can be changed by changing only its magnitude, or by changing only its direction, or by changing both its direction and its magnitude. A change of the velocity with no change in the magnitude is illustrated in Fig. 55. Let AB represent in magnitude and direction a certain velocity. A velocity BC is added to AB, giving as the resultant velocity AC. In the figure the line AC is the same length as the line AB. The magnitude of the velocity AB is not changed by adding BC, but its direction is. The velocity AC is a different velocity from AB.

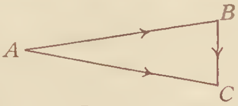


FIG. 55

Velocity is always a *vector* quantity, and vector methods must be used in adding velocities.

57. Projectiles. When a baseball is traveling through the air, it has at all times an acceleration toward the ground equal to g, where g is approximately equal to 32 feet per second per second, or 980 centimeters per second per second. It is a fact of fundamental importance that this acceleration is independent of any other motion the ball may have.

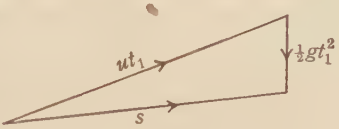


FIG. 56

The distance and direction of a projectile at any time t_1, measured from the time of starting, can be found graphically as follows: Draw the vector ut_1 (Fig. 56). This represents the distance the projectile would have traveled if the velocity had remained equal to the initial velocity u. To this vector add the vertical vector $\frac{1}{2}gt_1^2$, which represents the distance the projectile falls in the time t_1 (see equation 12). The resultant vector, s, gives the direction and the distance of the projectile from the starting-point.

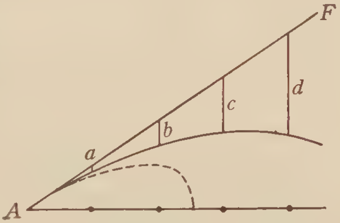


FIG. 57

By a combination of a number of figures like Fig. 56 the complete path of a projectile may be shown (Fig. 57). The line AF represents the path the body would take if it did not fall. The lines a, b, c, and d show the amount of fall in one, two, three, and four units of time. The curved line shows the line of flight. The actual path differs from this path on account of air friction. For very high speeds it would be modified a great deal, in the manner shown by the broken line of the figure.

It is instructive to use velocities as the vectors instead of distances and thus to find graphically the velocity at different instants during the flight. Figs. 58–61 show the result of the vector addition of the initial velocity and the downward velocity acquired in periods equal to t_1, t_2, t_3, and t_4 seconds. The quantities v_1, v_2, v_3, and v_4 are the vectors giving the direction and magnitude of the velocities at those instants. Obviously the value of the velocity at any desired time can be found from drawings made to scale.

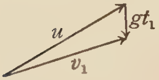


FIG. 58

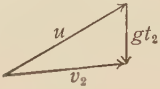


FIG. 59

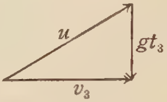


FIG. 60

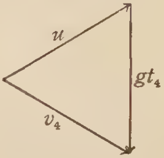


FIG. 61

The student has probably noticed that in adding these velocities equation (15) of section 55 was used, but not in the ordinary algebraic way. In this case the velocities are not in the same direction, and vector methods must be used.

An alternative method of attacking the problem of a projectile is based on the following considerations: The initial velocity u of the projectile may be divided into the two components shown in Fig. 62. The vertical component, u', is continually changed by the acceleration g, which acts in the vertical direction. On the other hand, the horizontal component, u'', is not affected by the downward acceleration. The result is that the horizontal component remains unchanged during the flight (except for air resistance), while the vertical component decreases to zero and reverses in direction.

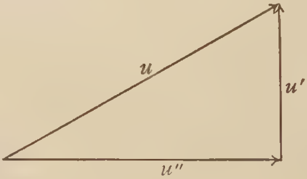


FIG. 62

If the numerical values of the horizontal and vertical components of the initial velocity are found from a graph like Fig. 62, it is not difficult to find the time of flight and the range, air resistance being neglected. The decrease in the vertical velocity each second is g. Hence the time required for the projectile to lose all its upward velocity is u'/g. Since this is the time required for the projectile to reach its highest point, it is equal to half the time of flight. Hence the time of flight, T, is given by

$$T = 2\frac{u'}{g}.$$

The horizontal distance traveled, or the range, is equal to the horizontal component multiplied by the time of flight, or

$$s = u''T.$$

If a table of sines and cosines is available, the following equations can be used to compute the values of the horizontal and vertical components:

$$u' = u \sin \theta,$$
$$u'' = u \cos \theta,$$

where θ is the angle the initial velocity makes with a horizontal direction. These equations follow from the definition of the sine and cosine of an angle.

When the angle is 45 degrees the two components are equal; and since the hypotenuse is $\sqrt{2}$ times one of the sides,

$$u' = u'' = \frac{u}{\sqrt{2}} = \frac{u}{1.41}.$$

The student may be interested to work out from these equations the range of an army rifle fired at an angle of 45 degrees with a velocity of 2700 feet per second. He will find that his result is many times larger than the actual range of such a gun. Air resistance for such high speeds is very important.

58. Independence of motions. In the last section it was assumed that the acceleration of a falling body is independent of any other motion that the body may have. At one time some thought that a rifle bullet did not begin falling until near the end of its flight. Experience shows that this is not true. The instant a ball leaves the gun it begins falling with an acceleration of 32 feet per second per second. If shot in a horizontal direction with a speed of 3000 feet per second, it will fall about 2 inches in 100 yards, and several feet in 500 yards. The sights of a rifle are usually adjusted to allow for a certain amount of fall.

On account of the motion of the earth we are carried through space with an enormous velocity. But this motion does not affect the playing of tennis, or baseball, or other activities. It is a general principle that the effect of any acceleration is independent of other motions that a body may have.

PROBLEMS

1. When a ship is sailing northeast at the rate of 10 mi./hr., with what velocity is it approaching a north-south coast which lies to the east?

2. A steamer is traveling north with a velocity of 20 mi./hr. The smoke from the funnel lies 30° east of south. If the wind is in the west, find its speed. (Use a graphical method.)

3. An automobile is moving at a speed of 30 mi./hr. up a grade which rises 10 ft. in each 100 ft. of roadway. Find the (*a*) horizontal and (*b*) vertical components of the velocity.

4. A man walks with a velocity of 5 mi./hr. across the deck of a steamer, in a direction making an angle of 45° with the velocity of the steamer. If the steamer's speed is 15 mi./hr., find, by the aid of a diagram, his resultant velocity.

5. An automobile starting from rest acquires a speed of 30 ft./sec. in 5 sec. Compute the average acceleration.

6. A sled, coasting downhill, has a uniform acceleration of 5 ft./sec.^{2} (*a*) What speed will it gain in 4 sec.? (*b*) How far will it go in 4 sec. if it starts from rest?

7. (*a*) How far will the sled of Problem 6 travel during the second second? (*b*) during the third and fourth seconds?

8. A ball starting from rest rolls down an incline with a constant acceleration equal to 10 cm./sec.^{2} (*a*) What speed will the ball have at the end of 5 sec.? (*b*) How far will it have rolled?

9. It required 20 min. for a train of two locomotives and ten passenger cars starting at rest to attain a speed of 60 mi./hr. Compute (*a*) the average acceleration and (*b*) the distance traveled while gaining this speed.

10. A heavy passenger train traveling 60 mi./hr. was stopped by the air brakes in 21 sec. Compute (*a*) the average acceleration and (*b*) the distance traveled while stopping.

11. A mass of 200 gm. falls from the top of a tower to the ground in 4 sec. (*a*) How high is the tower? (*b*) How fast will the mass be going when it strikes the ground? ($g = 980$ cm./sec.^{2} Neglect air resistance.)

12. A ball is thrown vertically with a speed of 80 ft./sec. (*a*) If air friction were negligible, how long would it be before the ball would reach its highest point? (*b*) How high would it go? ($g = 32$ ft./sec.^{2})

13. A stone is thrown vertically downward from a cliff with a speed of 20 ft./sec. In 3 sec. it strikes the ground below. (*a*) How high is the cliff? (*b*) With what speed will the stone strike? (Neglect air resistance. Assume $g = 32$ ft./sec.^{2})

14. A 5-ounce baseball was dropped 557 ft. from a window of the Washington Monument and was successfully caught. Neglecting air effects, (*a*) what was its final velocity? (*b*) How long did it take to fall?

15. If a baseball is thrown at an angle of 45° with the horizontal, with a speed of 122 ft./sec., how far will it travel? (Neglect air friction.)

16. A bullet is fired horizontally with a speed of 1000 ft./sec. How much will it fall while traveling a horizontal distance of 200 yd.? (All friction is neglected.)

17. A ball is thrown horizontally with a speed of 50 ft./sec. from the top of a tower 150 ft. high. How far from the base of the tower will it strike the ground?

18. A baseball is thrown horizontally from a high tower with an initial speed of 64 ft./sec. Show by vectors how to compute the direction and speed of the ball at the end of 2 sec.

CHAPTER VI

EFFECTS OF FORCES ON MOTION

Newton's first law of motion, 59. Force proportional to acceleration, 60. Mass, 61. Equal accelerations produced in different masses, 62. Equal forces acting on different masses, 63. The general relation between force, mass, and acceleration, 64. Units of force: the dyne; the pound, 65. Relationship between mass and weight, 66. Examples in the B.E. and C.G.S. systems, 67. Remarks on mass and weight, 68. Momentum; impulse, 69. Newton's first law of motion, 70. Newton's second law of motion, 71. Newton's third law of motion: action and reaction, 72. Conservation of momentum, 73. Motion in a curved path; centripetal force, 74. Magnitude of the centripetal force, 75. Centrifugal force, 76. Derivation of the formula for centripetal and centrifugal forces, 77. Applications of the foregoing principles, 78. Summary, 79.

59. Newton's first law of motion. What would happen to an automobile at rest if no forces acted on it? All of us know that it would stay at rest. But if the car were in motion and if no forces acted on it, what would happen? Before we answer this we should recall the fact that there are usually many different forces acting on an automobile in motion; among them forces of the nature of friction, such as air resistance, friction at the axles, and friction between the wheels and the ground. If we could eliminate *all* these forces, what would happen? This cannot be answered directly from experience, for in no case are we able to eliminate entirely all such forces; but it is a matter of common observation that as frictional forces are made less and less, there is a tendency in a moving body to run longer and longer. All experience points to this conclusion: *If all forces were eliminated from a moving body, it would continue to move with a constant velocity.*

We always find that it requires the constant application of a force to keep a wagon or a car in uniform motion on a level road, but this is because there are retarding forces of friction which oppose the motion. The important thing to know is this: that if the applied forces are *equal* and *opposite in direction* to the retarding forces, the car will continue to move with uniform velocity. In that case the *resultant* force acting on the car is zero. A larger driving force is necessary to keep an automobile running at a higher speed, because the retarding forces are greater for higher speeds. But the driving force is never greater than the retarding forces so long as the velocity is constant. In all cases where the motion is uniform and in a straight line the applied force is exactly equal to the retarding forces, and the resultant of all acting forces is zero.

We can now understand Newton's first law of motion, which may be stated as follows: *When the resultant of all forces acting on any body is zero, if at rest the body will remain at rest, if in motion it will continue in motion with a uniform speed in a straight line.*

This statement can be summed up in one word, *inertia.* On account of its inertia a body requires a force to put it in motion, but once in motion a force is necessary to change its velocity.

This law of Newton's is partly an assumption and partly the result of experiment. It is an assumption in the same sense that the axioms of geometry are assumptions. The strongest argument in favor of it is that conclusions based on it always agree with observation; no exception has ever been found. This law applies to all kinds of bodies: to the atom and molecule as well as to larger bodies, to those far out in space as well as to terrestrial ones.

The concept of acceleration is very important in connection with the action of forces. A body which is "at rest" or "in motion with uniform speed in a straight line" is one which has no acceleration. Hence Newton's first law may be stated as follows: *If the resultant force is zero, there will be no acceleration.*

Whenever there is an acceleration, there must be an *unbalanced* force, and the direction of this force is the direction of the acceleration. It will not be difficult for the student to find many illustrations of this fact. Since the passengers in a car are accelerated when the motion of the car is changed, there must be forces acting on them. We ride with comfort in trains or cars as long as the motion is uniform; but if there are irregularities (that is, accelerations), we are jerked first one way and then the other. The motion of the earth is so constant that the high speed brings no discomfort.

60. Force proportional to acceleration. This question at once arises: Is there any simple relation between the force acting on a car and the acceleration produced? For example, if the resultant force acting on the car be made twice as large, will the acceleration be twice as large? The answer is Yes. Experience shows that the acceleration is *directly proportional* to the resultant force. We may state this result of experience in the form of an equation and say that, for any one body,

$$\frac{F_1}{F_2} = \frac{a_1}{a_2}, \tag{17}$$

where a_1 is the acceleration produced by the force F_1, and a_2 the acceleration produced at another time by a force F_2. Both F_1 and F_2 are *resultant* forces acting on the body.

As an example the weight of any body is a force which will give it a downward acceleration of approximately 32 feet per second per second. To make this same body have an acceleration of 64 feet per second per second, or twice as much, the force must be twice the weight of the body. To give it an acceleration of 16 feet per second per second the resultant force must be half the weight. If the frictional force acting on a body sliding down vertical ways is equal to one fourth of the weight of the body, the resultant downward force will be three fourths of the weight, and the downward acceleration will be three fourths of 32, or 24 feet per second per second.

Equation (17) is the result of experience and is therefore called an experimental law.* The reason that quantitative laws are so commonly stated in the form of equations is that this method gives statements which are simpler and more definite than those expressed in words. But, on the other hand, the student will find that it will always add to his understanding of such an equation to translate the meaning into nonmathematical terms. Then the equation will mean something more than a mere collection of symbols.

---

*This statement needs some qualification. Man undoubtedly obtained his first concept of force from the weight of bodies. It is in this sense that the term has been used in the opening chapters of this book. When force is defined and measured in terms of weight, equation (17) is an experimental law. However, some authorities prefer to define force by equation (17) rather than by the weight of bodies. When force is defined in this manner, equation (17) cannot be regarded as an experimental law. In this case it becomes necessary to show by experiment that this concept of force agrees with that defined and measured by gravitational effects.

61. Mass. The word *weight* is the name given to the attractive force exerted by the earth. This force acts on objects around us because they happen to be near the earth. If this book were taken to the top of a mountain, it would weigh less. If it could be taken far out into space, the weight would become insignificant. At the center of the earth it would be attracted equally in all directions, and the resultant force would be zero. The weight is not something inherent in a body itself, but is an "accident of its environment."

The *mass* of a body is the same wherever it may be. A piece of lead placed at the center of the earth would have no weight; but the quantity of lead (the mass) would remain the same. Mass may be said to be another name for matter; it is used whenever the quantity of matter is to be measured.

In commerce it is usually not necessary to make a distinction between mass and the weight of the mass, but in scientific work many cases arise where a distinction is necessary. A method of comparing or measuring masses is indicated in the next section.

62. Equal accelerations produced in different masses. If a man kicks a barrel lying on its side, and then another barrel, he may observe that the acceleration of the first barrel is much greater than that of the second. If so, he concludes that the quantity of matter included in the two barrels (that is, the mass) is different. To have produced the same acceleration in each, a greater force would have been necessary for the barrel of greater mass. Another way to state this is to say that large masses have greater inertia than smaller ones. Why is it rougher riding over a rough road in a light automobile than in a heavy one? A light local train does not run as smoothly as the heavy through trains. A heavy Pullman on the rear of a light train will steady the motion of the entire train.

The quantitative statement is as follows: *To produce the same acceleration in two different masses the forces must be proportional to the masses*; or, stated algebraically, to produce the same acceleration,

$$\frac{F_1}{F_2}=\frac{m_1}{m_2}, \tag{18}$$

where F_1 is the force applied to the mass m_1, and F_2 the force applied to m_2. As before, F_1 and F_2 are resultant forces.

Everyone knows that a greater force is needed to throw (that is, to accelerate) a large stone than a small one. But this equation states more than this, for it states *how much* larger the force must be to give the same acceleration.

It should be noticed that forces are proportional to masses and not to weight. Far away from the earth bodies have practically no weight, yet it is believed that the same force is required to give a certain acceleration to a body out in space as would be necessary if that body were here on earth. This is assumed by astronomers in making computations of the masses and the accelerations of the moon and planets. The results confirm the validity of the assumption.

If the man had *lifted* the two barrels at a uniform speed, he would have compared their *weights*; but when he applied horizontal forces to them and observed the accelerations, he was comparing *masses*. Imagine this experiment to be repeated far out in space where the barrels would have no weight. There no force would be required to "lift" them, or move them away from the earth at a uniform speed; but the forces required to accelerate them would be the same as on the earth.

Equation (18) gives a method of comparing or measuring masses.

63. Equal forces acting on different masses. Imagine two small toy carts placed on a smooth table and connected by a stretched spiral spring. When released the two carts will be accelerated toward each other. The force acting on one is equal to that acting on the other, for each force is equal to the tension of the spring. If the masses are equal, the accelerations will be equal; but if a large mass is placed in one of the carts, its acceleration will be less than that of the other. This simple experiment does not admit of accurate measurements; but if it did, we should find what more accurate experiments show—that the acceleration of each cart *varies inversely as its mass.* This may be summarized by saying that when equal forces act on different masses,

$$\frac{a_1}{a_2}=\frac{m_2}{m_1}, \tag{19}$$

where a_1 is the acceleration given to mass m_1, and a_2 that given to mass m_2. This law can be stated as follows: *When equal forces act on two different masses, the accelerations produced are inversely proportional to the masses.*

Equations (17)–(19) are not independent. If any two of them are found to be true, the third can be derived from the others by a rather involved algebraic process. These three equations connect the fundamental concepts of mass, force, and motion (acceleration), and are therefore *fundamental* equations.

64. The general relation between force, mass, and acceleration. The three relations expressed by equations (17)–(19) can be included in one expression,

$$F \propto ma,$$

or

$$F = kma, \tag{20}$$

where k is a constant inserted in order to change the proportionality sign to one of equality. The student should *study this equation* and *see for himself* that it does include the three relations. When he understands that, he will understand why it is called the *general relation.* It should be kept in mind that the force F is always the resultant force.

Both the force F and the acceleration a in equation (20) are vector quantities and always have identical directions; in other words, the acceleration is always in the same direction as the force.

65. Units of force: the dyne; the pound. In order to understand the significance of the factor k in equation (20), a discussion of a well-known relation will be given. If A is the area, l the length, and b the breadth of a rectangular field,

$$A = kbl.$$

The numerical value of the factor k will depend on the units used in measuring the other quantities. If A is measured in acres, and b and l in rods, k will be equal to 1/160; but if b and l are measured in feet, k will have a different value. It is always possible to choose the units so that the factor k in such an equation is equal to unity. For example, if A is measured in square feet, and b and l in feet, k will equal unity. Hence the equation can be written

$$A = bl.$$

In a similar way the numerical value of the factor k in equation (20) will depend on what units are selected. The C. G. S. unit of force, called the *dyne*, is chosen of such size that the factor k is equal to unity when the mass m is measured in grams and the acceleration a in centimeters per second per second. The equation then becomes

$$F = ma. \tag{21}$$

By the aid of this equation it is a simple matter to compute the value of the force which is necessary to give a known mass a known acceleration; for example, the force required to give a mass of 500 grams an acceleration of 10 centimeters per second per second is 5000 dynes.

A convenient definition of the dyne can be obtained from equation (21) by taking the special case where both m and a are unity: *1 dyne will give to a mass of 1 gram an acceleration of 1 centimeter per second per second.*

In the B. E. system the unit of force, the *pound*, is determined independently of equation (21), for it is defined as the weight of a certain body. The unit of *mass*, however, is selected to be of such a value that the constant k in equation (20) will be unity. Hence equation (21) is true when the force is measured in pounds and the acceleration in feet and seconds. The name *slug* has been suggested for this unit of mass, but the suggestion has not been generally followed.

There is an important difference between the two systems of units. In the C.G.S. system, *mass* is the fundamental quantity, and the unit of force is defined by the aid of equation (21); while in the B.E. system, *force* is the fundamental quantity, and the unit of mass is defined by the aid of equation (21).

In the *British Absolute system* the unit of *mass* is called the *pound*, and the unit of force the *poundal*; but this system is rarely used outside of textbooks and will not be used in this book.

66. Relationship between mass and weight. There is, for all systems of units, a numerical relationship between the mass of a body and its weight. In the case of a freely falling body the force, F, is equal to the weight, and the acceleration produced by this force is g. Hence equation (21) becomes

$$\text{Weight of body} = \text{mass of same body} \times g. \tag{22}$$

In the C. G. S. system, force is measured in dynes, mass in grams, and $g = 980$ centimeters per second per second (approximately). Hence

$$\text{Weight in dynes} = 980 \times \text{mass in grams}. \tag{23}$$

By taking the special case where the mass is unity, it follows from this equation that 1 gram weighs 980 dynes. This is easily seen to be true by a different method of reasoning. In section 65 a dyne was defined as the force which will give to 1 gram an acceleration of 1 centimeter per second per second. To make a gram fall with an acceleration 980 times greater than this would require 980 dynes. Since weight gives it this acceleration, the weight of 1 gram must be 980 dynes.

In the B. E. system the weight is measured in pounds, and g* is equal to 32 feet per second per second (approximately).

$$\text{Weight in pounds} = 32 \times \text{mass},$$

or

$$\text{Mass} = \frac{\text{weight in pounds}}{32}. \tag{24}$$

As has been stated before, the name *slug* is sometimes used for the unit of mass in the B. E. system. Using this term, the mass which weighs 1 pound is by equation (24) equal to 1/32 of a slug.

67. Examples in the B. E. and C. G. S. systems. In both the B.E. and C.G.S. systems, equation (21),

$$F = ma,$$

holds for all cases when the proper units are used.

---

*The most probable value of g at sea level and latitude 45 degrees is 32.174 feet per second per second, and this number should be used where greater precision is desired.

If a force of 500 dynes is applied to a mass of 1 kilogram, it is easy to compute the rate of change in its velocity, or the acceleration. Substituting in equation (21),

$$500 = 1000 \times a,$$

or

$$a = 0.5 \text{ cm./sec.}^2$$

Take the case of a flexible cord hanging over a very light, low-friction pulley, with a kilogram hanging from each end of the cord.* When equilibrium is destroyed by adding an extra 100 grams to one side, the masses at once begin moving. The problem is to compute the acceleration. The force that causes the motion is the *weight* of the 100 grams, or 100×980 dynes. The total mass accelerated is $1000 + 1000 + 100$ grams. Hence

$$100 \times 980 = (1000 + 1000 + 100)a,$$

from which a can be readily computed. The mass and friction of the pulley have been neglected.

Find the acceleration that a force of 5 pounds would give to a body weighing 15 pounds. The fundamental relationship is

$$F = ma.$$

But in the B. E. system

$$m = \frac{\text{weight}}{32} = \frac{15}{32}.$$

Hence, substituting in equation (21),

$$5 = \frac{15}{32} \times a,$$

or

$$a = 10.7 \text{ ft./sec.}^2$$

Possible troubles when using the B.E. units may be avoided by writing equation (21) in the form

$$F = \frac{\text{weight}}{32} \times a. \qquad (25)$$

In some engineering treatises, in all formulas where *mass* usually appears, *weight/32* is substituted.

68. Remarks on mass and weight. On account of the proportionality between *mass* and *weight*, most persons do not distinguish between them. For example, in buying and selling commodities it usually makes no difference whether one refers to the mass or to the weight of the commodity he is handling. However, a person can, in most cases, determine whether it is the mass or the weight in which he is interested. When one buys sugar, he is really interested in the mass of it; he does not care whether the earth attracts it or not. On the other hand, if he buys stone for the ballast of a ship, he wants weight; he is then interested in forces.

---

* In the laboratory this apparatus is called Atwood's machine.

Ordinary "weighing" may be regarded either as a comparison of the *mass* of the substance with the *mass* of the balancing "weights" or as a comparison of the *weight* of the substance with the *weight* of the balancing load. It will tend toward clearer ideas if the student will form the habit of regarding weighing as a comparison of masses when he is using grams and as a comparison of forces when he is using pounds. The point is that he should fully and continually appreciate the fact that in the C.G.S. system the *gram* is a unit of *mass*, and that in the B.E. system the *pound* is a unit of *force*.

It is often convenient to use as an arbitrary unit of force the weight of a gram or a kilogram. Although the gram and kilogram are units of mass, their weights can be used as units of force. It is a good rule to say *gram* and *kilogram* when referring to *masses*, and *gram-weight* or *kilogram-weight* when referring to *forces*; for example, if one hangs a kilogram on the end of a vertical wire, he may say that the stretching force is 1 kilogram-weight, but not a force of 1 kilogram.

69. Momentum; impulse. It is a common expression to say that a moving object is difficult to stop on account of its *momentum*. The greater the momentum, the greater the effort required to stop it. A large and massive car has a larger momentum than a small one moving at the same speed. It is also true that the greater the velocity of the car, the greater is the momentum. Momentum thus depends on two factors: mass and velocity. The technical definition is as follows: *Momentum is equal to the product of the mass and velocity.* This is conveniently stated by the equation

$$\text{Momentum} = mv. \tag{26}$$

This definition contains the term *mass* and not *weight*. A moving body far out in space, where its weight would be nothing, has momentum. The earth itself has momentum, because it has mass and velocity (the expression "the weight of the earth" has no meaning).

In using the B.E. system it is convenient to substitute for mass its equivalent from equation (24). The defining equation for momentum is then in the form

$$\text{Momentum} = \frac{\text{weight}}{g} \times v. \tag{27}$$

For places where $g = 32$ feet per second per second this becomes

$$\text{Momentum} = \frac{\text{pounds weight}}{32} \times v.$$

No name for the unit of momentum is used in either the C.G.S or the B.E. system.

An *impulse* is defined as the *product of force times the time the force acts.*

Theorem. *The change of momentum produced by an impulse is numerically equal to that impulse.*

This will be illustrated by a numerical example: A force of 500 dynes acts on a mass of 1000 grams for 4 seconds. It follows from equation (21) that

$$\begin{aligned} a &= \tfrac{500}{1000} \\ &= 0.5 \text{ cm./sec.}^2 \end{aligned}$$

Hence the mass will have a uniform acceleration of 0.5 centimeters per second per second. At the end of 4 seconds it will have acquired a velocity of

$$0.5 \times 4 = 2 \text{ cm./sec.}$$

The change in momentum will be the mass times the change in velocity, or

$$1000 \times 2 = 2000.$$

This is numerically equal to the product of the force by the time,—that is, the impulse,— as demanded by the theorem.

Proof of the theorem. If a mass m has a velocity v_1, and a constant force acts on it for a time t and changes the velocity to v_2, the acceleration is

$$a = \frac{v_2 - v_1}{t};$$

and the force, from equation (21), is

$$F = m\frac{v_2 - v_1}{t}.$$

By a simple change this becomes

$$Ft = mv_2 - mv_1; \tag{28}$$

that is, the product of the force by the time it acts (the impulse) is equal to the change in the momentum.

70. Newton's first law of motion. Newton's first law (sect. 59) may be restated in another form: *The momentum of any body will remain constant unless the body has an unbalanced force acting on it.*

71. Newton's second law of motion. The first law states that with no force acting there will be no change in the momentum of a body. The second law goes farther than this and tells how much force is necessary to make a certain change in the momentum. It may be stated as follows: *The rate of change of momentum of a body is proportional to the impressed force and is in the same direction as that force.*

This law was anticipated when we derived equation (28), for this equation is a statement of the same principle. Equation (28) may be written as follows:

$$\text{Force} = \frac{\text{change in momentum}}{\text{time the force acts}}.$$

This equation states that the force is equal to the rate of change of momentum, while Newton's second law states merely that it is proportional. It is equal in this equation, because the units used here were so chosen that the factor k in equation (20) was equal to unity.

A formal statement of Newton's first and second laws adds very little to the completeness of the treatment given in this book, for they are both included in the law $F = ma$; but they are of considerable historic interest. They were stated by Sir Isaac Newton about 1687.

72. Newton's third law of motion: action and reaction. Newton's third law is as follows: *Every action has an equal and opposite reaction.*

The words *action* and *reaction*, as used above, mean *force*. This law, like the others, is the result of experience. We cannot push on anything without its pushing back. When we throw a ball, it pushes back on the hand. If anyone is in doubt about the ball's pushing back, he should go through the motion of throwing a ball, both with it in his hand and without. When a man is walking, there are two forces involved: he pushes on the ground, and the ground pushes on him with an equal force. When the ground is slippery, he finds it difficult to walk. Why? When one steps out of a boat, he goes in one direction, and the boat tends to go in the other. The driving-wheels of an automobile push on the road, and the road pushes back on the wheels with an equal force (Fig. 63). This book is attracted by the earth, and it attracts the earth with an equal force.

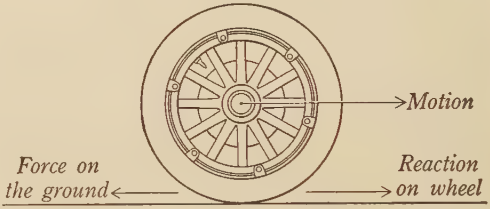


Fig. 63

In general this law states that forces always exist in pairs; each force has its equal and opposite mate. It is important to understand that action and reaction *never act on the same body*. Whenever a body *A* exerts a force on a body *B*, then *B* will exert an equal force on *A*. Consider the case of a book lying on a table. The earth attracts the book, and the table pushes up on the book. Since both these forces are acting on the book, they cannot be the action and reaction referred to in this law. The reaction to the attraction of the book is the attraction the book has for the earth. If we take the force the book exerts on the table as the action, the reaction is the force the table exerts on the book. To repeat: If *A* exerts a force on *B*, then the reaction is the force that *B* exerts on *A*.

73. Conservation of momentum. When a person steps out of a boat, he exerts a force on the boat, and the boat exerts an equal one on him. If no other forces act, the boat and the person will be given momenta which are opposite in direction and exactly equal. Stating this in the form of an equation:

$$m_1v_1 = m_2v_2, \tag{29}$$

where m_1 and m_2 are the masses of the person and the boat, and v_1 and v_2 the velocities acquired. It is a matter of common experience that in stepping out of a loaded rowboat very little velocity is given the boat; but the last person out must be careful, otherwise he may not step on the dock. The fact that the smaller the mass of the boat the greater will be its velocity follows from equation (29). The general law may be stated: *If a force and its reaction act between two bodies, and no other forces are present, equal and opposite momenta will be given to the two.*

This law can be derived by means of the theorem of section 69. Since the two forces, the action and its reaction, are equal, and the time of acting is the same for both, the impulses acting on the two bodies are equal; therefore the changes in momenta are equal.

Consider a group of bodies isolated from other bodies by an imaginary boundary. If no forces act across this boundary, the changes in momenta produced by forces acting between any two bodies of the group will always be equal and in opposite directions. Hence the change in the total momentum of the group will be zero. An automobile, or a train, and its contents may be regarded as such a group. An occupant of an automobile cannot change the total momentum of the car and its contents by any force that he may exert. Internal forces alone cannot stop a car or a train. The stopping force is the friction between the tire and the ground or between the wheel and the rail. Anyone who has attempted to stop an automobile when the ground was covered with smooth ice knows this very well.

Only external forces can change the momentum of a body or of a group of bodies.

The fact that a group of bodies tends to maintain its momentum unchanged is often referred to as the *conservation of momentum.*

74. Motion in a curved path; centripetal force. Every boy knows that when he whirls a stone around on the end of a string, there is danger of breaking the string. He also observes that the faster he makes the stone go, the harder must he pull on the string. In other words, he observes that in order to hold the stone in this circular path a force acting toward the center is required, and that the greater the speed of the stone, the greater must this "central" force be. This is found to be true in all cases where anything is being moved in a curved path. Mud and water may be thrown off a revolving wheel when there is not sufficient adhesive force to hold them in the circular path. Forces are required to hold the different parts of a flywheel in the circular path, and if it is not strongly built, at high speeds it may fly to pieces.

From these and many other similar experiments, we may derive the generalization that since the natural tendency of a body is to move in a straight line, when it is moved in a curved path forces are required (this statement is a corollary of Newton's first law). When the body is moved in a circular path, the force which must act toward the center of the circle is called the *centripetal* force or, more simply, the *central* force.

75. Magnitude of the centripetal force. It is not sufficient to know that a force is required in order to constrain a body to move in a curved path. Many cases arise where it is necessary to know how to compute the magnitude of the force. The formula for computing the magnitude, together with an explanation of its use, will be given, and, later, its derivation.

The force required to hold a mass m in a circular path of radius r while it is moving with a speed v is given by the equation

$$\text{Centripetal, or central, force} = m\frac{v^2}{r}. \qquad (30)$$

If the C.G.S. system is used, the force will be in dynes, the other quantities being measured in grams, centimeters, and seconds.

When the B.E. system is used, the weight in pounds of the moving body will probably be given instead of its mass. In that case *weight/32* is substituted for the mass m of equation (30). Then the force required to hold the body in a curved path is

$$\text{Force in pounds} = \frac{\text{weight in pounds}}{32} \cdot \frac{v^2}{r} = \frac{wv^2}{32\,r}, \qquad (31)$$

where v is measured in feet per second, and the radius in feet.

As an example a computation will be made of the tension on a string 3 feet long, by means of which a body weighing 1 pound is whirled in a circular path at an angular speed of 2 revolutions per second. The linear speed of the body (that is, the distance traveled in 1 second) is

$$v = 2\pi \times 3 \times 2 = 12\pi \text{ ft./sec.}$$

The tension on the string, or the force required to hold this weight in the circular path, is given by equation (31). Substituting the numerical values, we have

$$\frac{1}{32}\,\frac{(12\pi)^2}{3} = 1.5\pi^2 = 14.8 \text{ lb. (approximately).}$$

If the body were whirled in a vertical circle, the tension would not be the same at all parts of the path; for the gravitational attraction of the earth, or the weight, must be taken into account. In many cases the central force is so much larger than the weight that the effect of the weight may be neglected.

76. Centrifugal force. Fig. 64 shows the circular path of a stone held by a string. F_1 represents the central force which the string exerts on the stone. To this force there must be an equal but opposite reaction —the pull of the stone on the string. This force is represented by F_2, which must be equal to F_1 but opposite to it. The name *centrifugal force* is given to the force F_2. It should be clearly understood that the centrifugal force F_2 *acts on the string and not on the stone.* When the string breaks, there will be no force acting on the stone. It will then fly off in the direction v, along a tangent, simply because there are no forces acting on it. Newton's first law tells us that bodies move in straight lines when no force acts on them. There is momentum in the direction of v, but no force.*

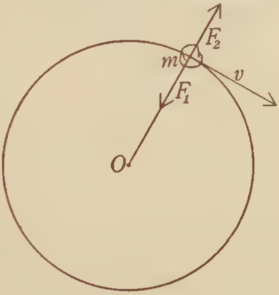


FIG. 64

---

* Two simple facts disprove the existence of a force in the direction of v: first, a force in the direction of v would not break the string; second, such a force would tend to make the stone move with an increasing speed around the circle, a tendency which we know does not exist.

To summarize: The *centripetal* force, F_1, holds the stone in its orbit; the *centrifugal* force, F_2, is the pull of the stone on the string. When the string breaks, the forces F_1 and F_2 disappear, and the stone flies off at a tangent. The two forces F_1 and F_2, being "action and reaction," are always equal and opposite.

When an automobile turns a corner, the friction between the tires and the ground is the force which holds the car in the curved path. The force exerted by the roadway on the car is the centripetal force, and the force, or reaction, which the car exerts on the roadway is the centrifugal force. In case the frictional force between the tires and the ground is not enough to swing the car around in the curved path (that is, is less than $wv^2/32\,r$ pounds), the car skids, refusing to make the turn.

77. Derivation of the formula for centripetal and centrifugal forces. Before deriving the formula for centripetal and centrifugal forces it is necessary to remind the student of one of the common methods of measuring an angle. According to this method the angle θ (Fig. 65) is given by the relation

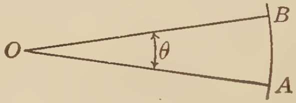


FIG. 65

$$\theta = \frac{\text{arc}}{\text{radius}}, \tag{32}$$

or, in terms of the symbols in the figure,

$$\theta = \frac{AB}{OA}.$$

When the angle is measured in this way, the numerical result is in *radians*. (1 radian is equal to 57.3 degrees. An angle of 360 degrees is equal to 2π radians, 180 degrees to π radians, and 90 degrees to $\pi/2$ radians.)

To derive the formula used in the last section it will be necessary to show that a body which moves in a circular path has an acceleration toward the center of the circle, and then to derive a formula for the numerical value of this quantity. After we have computed the value of the acceleration, it is very easy to find the force, for it can be found by the law $F = ma$.

Let the vector v in Fig. 66 represent the velocity of a body revolving at constant speed about the axis O. When the body reaches the point N, the velocity v_2 has the same magnitude but a different direction. In Fig. 67, AB represents the velocity at the point M of Fig. 66, and AC the velocity when the body is at N. From the methods of vector addition it is seen that the velocity which must be added to AB to produce AC is BC, or v'. If the interval of time required for the body to go from M to N is τ, the average rate of change of the velocity, or the acceleration, is

$$a = \frac{v'}{\tau}. \tag{33}$$

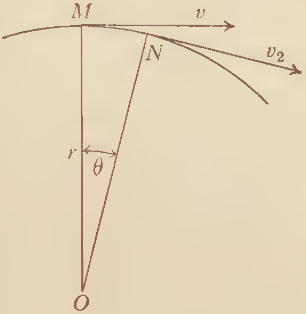


FIG. 66

In Fig. 66 the angle MON, or θ, is equal to the angle BAC of Fig. 67, for the lines AB and AC in Fig. 67 are parallel to the tangents at M and N. The position of the two points M and N must be so chosen that the angle θ is small (it is exaggerated for clearness in the figure). From Fig. 66 (equation (32))

$$\theta = \frac{\text{arc } MN}{r}.$$

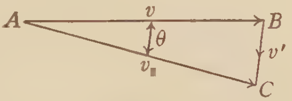


FIG. 67

But the arc MN is equal to $v\tau$, where τ is the time it takes for the body to move from M to N. Hence

$$\theta = \frac{v\tau}{r}. \tag{34}$$

In Fig. 67
$$\theta = \frac{\text{arc}}{\text{radius}} = \frac{BC}{AB};$$

for the angle is so small that the line BC may be taken as equal to the arc drawn from B to C, with A as its center. Since BC is equal to v', which, by equation (33), is equal to $a\tau$, and since AB is equal to v,

$$\theta = \frac{BC}{AB} = \frac{a\tau}{v}. \tag{35}$$

Equating the two values of θ given by equations (34) and (35),

$$\frac{v\tau}{r}=\frac{a\tau}{v};$$

and solving for the acceleration *a*,

$$a=\frac{v^2}{r}. \tag{36}$$

This equation gives the value of the acceleration of a body moving in a circular path of radius r with a speed v.

Since θ is very small, v' is perpendicular to v (Fig. 67); hence the velocity to be added is always at right angles to the velocity v. If v' is perpendicular to v, then a, which is the rate at which v' is added, must also be perpendicular to v. Hence *the acceleration is always directed toward the center of the circle.*

Whenever the acceleration and the mass are known, the force can be computed. In this case

$$\text{Force}=ma=\frac{mv^2}{r}. \tag{37}$$

This gives the value of the centripetal force (also the centrifugal force, for the two are always equal). The centripetal force is always in the same direction as the acceleration it produces. Hence, since the acceleration is toward the center, *the force required to hold the mass in a circular path is always directed toward the center of the circle.*

A simple relation between the linear speed v and the angular speed in revolutions per second, n, enables equation (37) to be written in a different form. If the value

$$v=2\pi rn$$

is substituted in equation (37), the force is given by the equation

$$F=4\pi^2n^2mr. \tag{38}$$

78. Applications of the foregoing principles. There are two different ways in which the principles discussed in the last sections have been applied. In one the momentum of a body acquired by rotating it is used to project it forward. For example, in that ancient weapon the sling, a stone was whirled around until it had acquired the desired momentum. When released, it would fly off at a tangent, with a large momentum. In the other method the centrifugal force (the force acting away from the center) is utilized. Examples of this are the cream separator, the centrifugal blower, the centrifugal pump, and the centrifugal drier. The centrifugal force, being proportional to the square of the speed, attains very large values in those machines which run at high speeds.

The cream separator. Cream is suspended throughout fresh milk in small globules, forming a mixture called an emulsion. These tiny spheres have a smaller specific gravity than the milk and, if left alone, will rise slowly to the top. They do not rise rapidly, for the same reason that small dust particles in the air do not fall rapidly. When "whole" milk is placed in a separator, and rotated at a very high rate of speed, the heavier milk goes to the outside and the lighter cream particles to the center. This may be explained as follows: The formula (equation (37) or (38)) states that centripetal force is proportional to the mass m. Hence, in order to hold a body in circular rotation, the centripetal force must be larger for bodies of larger mass. Larger forces are required to hold the milk in rotation than are required for the lighter cream particles. But the pressure at any one place in the liquid must be uniform. If it is large enough for the milk, it will be too large for the cream, and the cream will be forced toward the center.

The centrifugal drier. The centrifugal drier is a large cylinder with perforated sides in which wet clothes are placed. When this is whirled at a high speed, the adhesive force is not sufficient to hold the water in the clothes; hence the water is thrown out through the openings in the sides of the cylinder.

The leaning bicycle rider. When a man is running at a high speed around a corner, he must lean toward the corner. If a bicycle rider goes around a curve, he too must lean. Why? In both these cases, as in all cases of motion in curved paths, there must be a central force to change the motion from a straight line to a curved one. The explanation which follows is for the case of the bicycle rider, but it can readily be adapted to other cases. Let the line AB be drawn, from the point where the bicycle touches the ground, up through the center of gravity O

of the rider and bicycle (Fig. 68). The vector *OW* represents the downward force, the weight. The vectors *OC* and *OD* are two components of the weight. There is another force acting, the one with which the ground pushes on the tire (not shown in the figure). This acts in the direction *AO* and is equal and opposite to the force *OC*, thus annulling the effect of *OC*. The resultant force acting on the rider is *OD*. It is the central, or centripetal, force that is used to hold the bicycle and rider in the curved path.

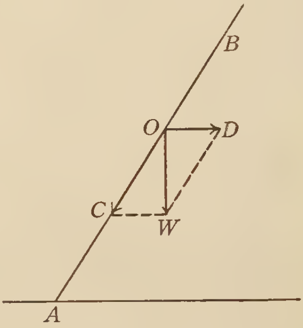


Fig. 68

The outer rail on a curve of a railroad track is made higher, in order to make the trains lean. A race track is "banked" at the curves for the same reason. On the ordinary street-crossing the turn cannot be banked; hence an automobile must depend on the friction between the road and the tires to furnish a central force.

79. Summary. This chapter contains some of the most important principles of physics, and the student should persist until he understands them clearly.

When different forces act on the same mass at different times, *the accelerations are proportional to the forces.*

When equal accelerations are produced in different masses, *the forces are proportional to the masses.*

When equal forces act on different masses, *the accelerations are inversely proportional to the masses.*

If the student will not attempt to commit these laws to memory but will imagine actual experiments showing the facts in question, he will have no difficulty in remembering the laws.

The above statements are included in formula (20),

$$F = kma,$$

which is a statement of one of the great principles of physics. In both the C.G.S. and B.E. systems the units are so chosen that $k = 1$. Hence (equation (21))

$$F = ma.$$

The *dyne* is the unbalanced force that will give 1 gram an acceleration of 1 centimeter per second per second. The weight of a gram is 980 dynes.

The weight of a mass m is mg, or (approximately) 980 m dynes. The B.E. unit of mass, the *slug*, is of such a size that

$$\text{Mass of a body} = \frac{\text{weight in pounds}}{g} = \frac{w}{32} \text{ slugs (approximately).}$$

It is convenient when using B.E. units to use formula (25),

$$F = \frac{\text{weight}}{32} a.$$

Momentum is equal to mass times velocity. It is a vector quantity having the same direction as the velocity.

Impulse is equal to *force times the time the force acts.* An impulse when unbalanced will produce a change in momentum numerically equal to the impulse.

Newton's first law states that when the resultant of all forces acting on a body is zero, if the body is at rest it will stay at rest, and if in motion it will continue in motion with a uniform speed in a straight line. Stated in other words: The momentum of a body will remain constant unless the body is acted upon by an unbalanced force.

Newton's *second law* may be stated: The rate of change of the momentum of a body is proportional to the impressed force and is in the same direction as that force.

The equation $F = ma$ covers both the first and second laws.

Newton's *third law* is as follows: Every force has an equal and opposite reaction. Forces always exist in pairs, but the two never act on the same body. If a body A exerts a force on a body B, then B exerts an equal and opposite force on A.

When A exerts a force and changes the momentum of B, there will be an equal and opposite change in the momentum of A unless other forces act on A to prevent it. When no external forces act on a group of bodies, their total momentum remains constant (conservation of momentum).

When a body is forced to move in a curved path, an acceleration must be given to it. If the body moves with a speed v in a path of radius r, the acceleration is v^2/r.

The centripetal, or central, force is mv^2/r; or, if the B.E. system is used, it is $\frac{wv^2}{gr}$, where w is the weight in *pounds* and g is approximately equal to 32 feet per second per second.

The centrifugal force acts away from the center, being always equal and opposite to the central force. These two never act on the same body. In the case of a stone whirling on the end of a string, the pull of the string on the stone is the central force, and the pull of the stone on the string is the centrifugal one.

PROBLEMS

1. (*a*) What is the acceleration when a force of 40 dynes acts upon a mass of 5 gm.? (*b*) How far will the mass move in 4 sec. if it starts from rest?

2. A body of mass 500 gm. is started sliding along a fairly smooth floor with an initial speed of 100 cm./sec. It is brought to rest by friction in 5 sec. Find (*a*) the value of the (negative) acceleration and (*b*) that of the force that brings the body to rest.

3. (*a*) What force (state the name of the unit as well as the numerical value) is required to give 200 gm. a speed of 200 cm./sec. in 2 sec.? (*b*) After it has acquired a speed of 240 cm./sec., what force would be required to stop it in 4 sec.?

4. Compute the force, in dynes, of the attraction of the earth for a mass of 20 gm.

5. Find the mass, in slugs, of a man who weighs 160 lb.

6. A certain body weighs 30 lb. on the earth. Suppose that this body were far away in space, and that a force of 15 lb. acted on it. What would happen?

7. What force in pounds will give to a body weighing 120 lb. an upward acceleration of 40 ft./sec.^{2}?

8. How many seconds would it take an engine pulling with a force of 2 tons more than friction to get a 640-ton train from rest up to a speed of 60 mi. an hour (88 ft./sec.) on a level track?

9. A boy and sled with a total weight of 96 lb. are pushed on very smooth ice by a horizontal force of 6 lb. (*a*) What speed is gained in 3 sec.? (*b*) What distance will be traversed in 5 sec.?

10. A horizontal force of 10 lb. acts on a baseball weighing 5 oz. for one fifth of a second. What velocity will be given the ball?

11. Two masses of 40 and 50 gm., respectively, are attached to the opposite ends of a perfectly flexible string which passes over a frictionless pulley, the mass of which may be neglected. (*a*) What force will be active in setting the two masses in motion? (*b*) What will be the acceleration?

12. A 150-pound man is hanging by a rope from a balloon which is falling with an acceleration of 4 ft./sec.^{2} What is the tension on the rope?

13. A man who weighs 150 lb. gives an impulse to a 75-pound canoe as he steps out of it. If the canoe is given a velocity of 2 ft./sec., compute the following: (*a*) the momentum given to the canoe, (*b*) the momentum the man acquires, and (*c*) the velocity the man acquires.

14. Two masses of 200 and 600 gm., respectively, move toward each other on smooth ice at the same speed, 20 cm./sec. (*a*) Compute the total momentum before collision. (*b*) If they stick together on impact, compute their common velocity. (*c*) Compute the numerical value of the impulse.

15. A 1000-pound animal is running toward a hunter at a speed of 10 ft./sec. A bullet weighing 1 oz. is fired, with a speed of 2000 ft./sec., into this animal. (*a*) Would the bullet have momentum enough to stop the animal? (*b*) If not, how much speed would the animal have left?

16. A horizontal force of 10 lb. acting for one fifth of a second is used to throw a stone weighing half a pound. (*a*) What is the numerical value of the impulse? (*b*) the momentum given the stone? (*c*) the velocity given the stone?

17. A mass of 1000 gm. is moving in a circular path of radius 20 cm. at a speed of 30 revolutions per minute. What force is required to hold it in its path?

18. What force is necessary to keep 600 gm. moving in a circle of radius 40 cm. at 120 revolutions per minute?

19. A 3000-pound automobile goes around a curve of 100-foot radius at 30 mi. per hour. Compute the side-thrust on the roadway.

20. Compute the centrifugal force exerted by a 150-pound man while he is being carried by an automobile around a curve of 50-foot radius at a speed of 20 mi. per hour.

CHAPTER VII

WORK AND ENERGY

Work, 80. Units of work, 81. Examples, 82. General method for computing work, 83. Energy, 84. Measurement of energy, 85. Two forms of energy, 86. Transformation of energy, 87. Method of computing kinetic energy, 88. Method of computing potential energy, 89. Power, 90. Conservation of energy, 91. Availability of energy, 92. Sources of energy, 93. Comparison of the C.G.S. and B.E. systems of units, 94. Summary, 95.

The concepts of work and energy are among the most important in the whole realm of physical science. The laws of energy developed in the subject of mechanics and originally applied only to bodies of finite size are now known to be far more general, applying not only to the smallest subdivisions of matter, such as the molecule, the atom, and the electron, but also to such things as electrical and magnetic fields and to the radiations of light-waves and electrical waves; in fact, so far as we know, to everything. Many of our common, daily observations can be clearly understood only when interpreted in terms of work and energy; therefore a clear understanding of these concepts is necessary for many persons and desirable for all.

80. Work. The following are simple cases of *work*. A man carries a hod of bricks up a ladder. Another shovels dirt out of a ditch. Another lifts stone into a wagon. In each of these cases *force* is exerted. In measuring the amount of work done the magnitude of this force must be taken into account. Another factor enters; namely, the *distance* through which the stone or dirt moves while acted upon by this force. If the man exerts force through a greater distance (for example, if he lifts the stone higher), more work is done. The man who digs a deeper ditch does more work in lifting a certain amount of dirt than one digging in a shallow ditch. Two factors always enter when one wishes to estimate the amount of work done: (1) the force and (2) the displacement, which is the distance the body upon which the force acts moves while the force is exerted.

By common agreement, work is defined as *the product of force exerted and the distance through which that force moves.*

When 10 pounds is used in lifting a weight from the floor to a table 3 feet high, the work done is measured by the product of 10 and 3. When work is measured by multiplying pounds and feet, the product is called *foot-pounds*. Thus, in this case, there are 30 *foot-pounds* of work done.

This method of computing work is more conveniently stated in the form of an equation:

$$\text{Work} = \text{force} \times \text{displacement in direction of the force},$$

or

$$W = Fs. \tag{39}$$

This definition of work restricts it to those cases where there is motion of some kind. Popularly the term is not always used in so limited a sense. Sometimes it is used to apply to cases where there is no force and no motion; for example, to mental work. It is also used in other cases in a vague, indefinite way. But a scientific definition of a physical quantity that can be measured must be both consistent and quantitative. Popular usage of the term *work* is not always consistent. For example, if a man standing still holds up the end of a wagon, he would say that he is doing work; but if the wagon is being held up by an inanimate prop, he would probably say that no work is being done, for it is customary to say that machines which are not moving do no work. In both cases forces are involved in a similar way, and the result of their application is the same. Why should there be work in one case and not in the other? For both these cases the technical definition is consistent, telling us that no work is done. because there is no motion in the direction of the force. No work is done if either the force or the motion becomes zero in any of the following cases: various types of engines, windmills, pushing a lawn-mower, sweeping, rowing, painting, sawing, hoeing, plowing, and so on.

81. Units of work. Many units are used for measuring work. The large number is due to the fact that the quantity is measured in so many different kinds of industry. Most of these units are based directly on equation (39).

In the B.E. system the force F is measured in pounds and the distance s in feet. The product is called *foot-pounds*. A *foot-pound* is the work done when a force of 1 pound is exerted on a body and there is a displacement of 1 foot in the direction of the force. Sometimes, when large forces are involved, the work is measured in *foot-tons*. One *foot-ton* is equal to 2000 *foot-pounds*.

In the C.G.S. system the force is measured in dynes and the distance s in centimeters. The product is called *ergs*.* An erg is the work done when a force of 1 dyne is exerted on a body and there is a displacement of 1 centimeter in the direction of the force.

The erg is such a small unit that a multiple of it, the *joule*, is often used. One joule is equal to 10,000,000, or 10^7, ergs.

Other units used are the *kilowatt-hour* and *horse-power hour*, in engineering work, and the *calorie* in thermal and chemical processes. The following numerical equivalents are given for reference:

1 joule = 10^7 ergs
1 joule = 0.738 foot-pounds
1 joule = 0.000000278 kilowatt-hours
1 joule = 0.000000373 horse-power hours
1 foot-pound = 1.35 joules
1 foot-pound = 1.35×10^7 ergs
1 calorie = 4.18 joules
1 calorie = 3.086 foot-pounds
1 kilogram-calorie = 1000 calories
1 kilowatt-hour = 3.6×10^6 joules
1 kilowatt-hour = 2.655×10^6 foot-pounds
1 horse-power hour = 1.98×10^6 foot-pounds
1 horse-power hour = 2.68×10^6 joules
1 horse-power hour = 0.746 kilowatt-hours
1 kilowatt-hour = 1.341 horse-power hours

82. Examples. The average pressure developed in the cylinder of a certain gas engine by an explosion is 80 pounds per square inch. If the area of the piston is 10 square inches and the length of the stroke is 8 inches, how much work is done by the explosion? The total average force acting on the piston is 80 × 10, or 800, pounds. The distance moved is 8 inches, or two thirds of a foot. The work in foot-pounds is found by multiplying 800 by 2/3.

---

*From a Greek word meaning "work."

A man weighing 150 pounds carries 100 pounds up a flight of stairs, a total vertical height of 12 feet. The work done is $(150+100) \times 12 = 3000$ foot-pounds. The force involved in this problem is a vertical force, and the distance used must be the vertical distance, not the distance the man actually goes.

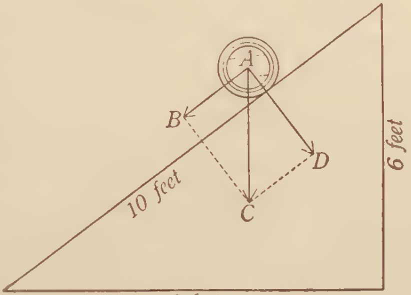


FIG. 69

A barrel that weighs 300 pounds is rolled up an inclined plane by a force acting parallel to the plane. The dimensions are indicated in Fig. 69. Before undertaking to compute the work done, the force (acting parallel to the plane) which is necessary to give the barrel uniform motion up the plane will be computed by the methods of section 7. The weight *AC* can be resolved into two components: *AB* parallel to the surface of the plane and *AD* perpendicular to it. When the friction is negligible, the force that will give the barrel a uniform motion up the plane will be equal and opposite to the component *AB*. Hence, if we find the value of the vector *AB* and multiply it by the length of the incline, we shall obtain the numerical value of the work done in rolling the barrel up the incline.

The triangle *ABC* is similar to the large triangle. Hence, from the corresponding sides,

$$\frac{AB}{AC}=\frac{6}{10}.$$

Since the force *AC* is 300, we have

$$AB = 300 \times \tfrac{6}{10} = 180 \text{ lb.}$$

Hence the work in rolling the barrel up the plane is

$$180 \times 10 = 1800 \text{ ft.-lb.}$$

The work done when the barrel is lifted directly up the height of the plane will now be computed. The force required is 300 pounds, and the distance is 6 feet. In this case a vertical force is used; to compute the work, a vertical distance must be taken.

$$\text{Work} = 300 \times 6 = 1800 \text{ ft.-lb.}$$

The same work is done in lifting the barrel straight up as when it is rolled up the incline, but in one case a much greater force is required.

In the preceding cases the force F and the distance s in the formula $W = Fs$ were parallel to each other. When the force parallel to the incline was used, the distance s was the length of the incline. When a vertical force was used, s was a vertical distance. Not only in these cases but *in all cases the force F and the distance s must be parallel to each other.*

To make this point clearer, suppose a weight of 50 pounds is lying on the table and is pushed along the table a distance of 3 feet. How shall we compute the work? It would not be correct to say that the work is 50 × 3. Why? Sufficient information is not given to solve this problem. The force acting in the direction of the 3 feet is unknown. It is the force necessary to overcome the friction between the weight and the table. The force of 50 pounds acts vertically, and there is no vertical motion; hence the force of 50 pounds does no work. To repeat: The force F and the distance s must be in the same direction.

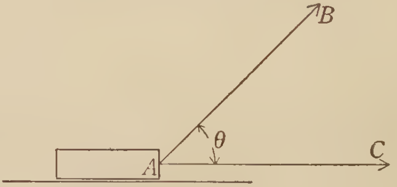


FIG. 70

83. General method for computing work. In Fig. 70 the force is represented by the vector AB, and the distance that the body is dragged by this force is represented by the vector AC. Either one of the two following methods may be used to compute the work done:

1. The work is equal to the force times the component of the displacement that is parallel to the force. The component of AC parallel to the force AB is obtained by drawing the line

DC perpendicular to AB (Fig. 71). The desired component is AD. Hence

$$\text{Work} = AB \times AD. \tag{40}$$

2. The work is equal to the component of the force parallel to the displacement times the distance AC. The component of the force is AE. Hence

$$\text{Work} = AE \times AC. \tag{41}$$

These two expressions for the work are always numerically equal; for it is easily proved, from the similar triangles of Fig. 71, that the right-hand sides of equations (40) and (41) are equal.

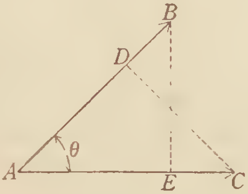


FIG. 71

If θ is the angle between the direction of the force and the direction of the motion, equations (40) and (41) are both equal to the formula

$$\text{Work} = AB \times AC \cos\theta. \tag{42}$$

For in the triangle ADC, by definition of the cosine,

$$\cos\theta = \frac{AD}{AC}, \quad \text{or} \quad AC\cos\theta = AD.$$

Substituting this in equation (42), equation (40) is obtained. In the triangle ABE, by definition,

$$\cos\theta = \frac{AE}{AB}, \quad \text{or} \quad AB = \frac{AE}{\cos\theta}.$$

Substituting this in equation (42), equation (41) is obtained.

Since AB represents the force F, and AC the distance moved, we can, from equation (42), say that, in general,

$$\text{Work} = Fs \times \cos\theta, \tag{43}$$

where θ is the angle between the force and the direction of the motion, and s is the actual distance the body moves.

Summary. Work is equal to the product of force times distance. But the force and distance must be parallel to each other. To get the correct force and distance, method 1 or 2 must be used.

84. Energy. Often when work is done on an object, something is given to the object which it retains and which later enables it to do work. When a clock spring is wound up, it acquires, through having work done on it, the ability to do work itself and is able to keep the clock running. That which has been stored in a body and enables it to do work is called *energy*. The clock spring, in being wound, gains energy; and when it does work in driving the clock, it loses this energy. The work done in lifting the weight of a pile-driver gives it energy. When it falls, it can do work. When a rifle is fired, the exploding gases do work on the bullet, giving it energy; when the bullet strikes an object, it does work and loses its energy.

In general, any body that is in such a condition that it is able to do work is said to have energy.

85. Measurement of energy. Today energy is a marketable commodity; it is commonly bought and sold. When one pays his electric-light bill, he pays for the amount of energy delivered to him; for electric meters are so designed that they indicate the amount of energy supplied. Many small factories and mills buy their energy from neighboring concerns. In order to buy and sell energy, methods must be devised for measuring it.

Changes in the energy of any body or any machine are always associated with work. A boy throwing a baseball does a certain amount of work on the ball and loses a certain amount of energy; the energy gained by the ball is equal to the work the boy does. (This assumes that there are no frictional losses, that the boy does no work on the air. If he does, then part of the work he does gives some energy to the air.) If a weight of 100 pounds is lifted 10 feet, 1000 foot-pounds of work have been done, and the weight is said to have gained 1000 foot-pounds of energy. In falling it can do 1000 foot-pounds of work.

The energy a body possesses is equal to the total work it can do.

Energy is not always gained by a body when work is done on it. A wagon drawn at a constant speed on a level road is not gaining or losing energy. In that case all the applied force is used in overcoming friction. The work done is used to develop heat and noise, which are forms of energy. When the wagon goes up a hill or increases its speed, it gains energy, but on account of friction the gain in energy is less than the work done. If 30 per cent of the work is lost on account of friction, the energy gained by the wagon is equal to 70 per cent of the total work done and may be said to be equal to the net amount of work done on the wagon.

In general, the loss in energy of a body may be measured by the work it does when losing energy, and the gain in energy of a body may be measured by the net amount of work done on it.

86. Two forms of energy. Bodies can possess energy in a great many different ways. In general there are two classes:

1. Bodies in *motion* have energy, for in losing their speed they can do work. That form of energy which is due to motion is called *kinetic energy*. A moving train or automobile, a ball or a bullet in flight, a rotating flywheel, have kinetic energy.

2. All other forms of energy are called *potential energy*. The more common cases of potential energy may be grouped under three heads:

(*a*) Potential energy may be due to the position of a body with respect to the ground. A stone on top of a building, water in an elevated reservoir, and the lifted weight of a pile-driver have potential energy because of their position.

(*b*) Potential energy may be due to the distortion of an elastic body. A stretched or compressed spring has potential energy.

(*c*) Potential energy may be possessed by those substances that can do work through some chemical action. Thus the energy of coal, gasoline, foods, etc. is potential energy of this type.

87. Transformation of energy. Simple observation reveals many cases where energy changes from one form to another. A boy sitting in a swing is pulled out to one side, and thus raised off the ground. The work done gives him potential energy. When released he swings back, losing his potential energy and gaining energy of motion, or kinetic energy. As he passes the lowest point and goes toward the other end, he loses kinetic energy and again gains potential energy. A vibrating clock pendulum is a similar case of the continuous changing of energy from one form to another.

In Fig. 72 are shown an axle and a flywheel. When the rope is wound around the axle, the heavy weight is lifted and gains potential energy. When it falls, this potential energy changes into kinetic energy of the flywheel and of the weight. Part of the energy is stored in the wheel. When the weight strikes the floor, the wheel and axle will keep on running, winding the rope up on the axle and lifting the weight. This storing of energy in a flywheel is very common. In a gas engine the energy is not supplied continuously, but at intervals. It is the energy stored in the flywheel that makes it possible to get from this engine a nearly uniform supply of energy. Many types of machines have flywheels which help maintain an approximately uniform speed.

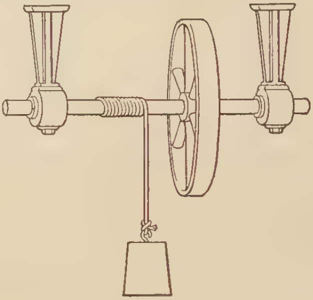


FIG. 72

A city electric-power plant affords an illustration of complicated transformations of energy. Part of the energy in coal is converted into energy of steam. By means of an engine this is changed into energy of moving machinery, then by the aid of a dynamo into electrical energy, which is transmitted over wires and, at some distant point, is changed into light, or heat, or energy of moving machines.

Since the temperature of a man is usually higher than his surroundings, he is continually losing energy in the form of heat. He loses energy, also, on account of the physical work he does. To compensate for these losses a moderately active man must receive daily about 10,000,000 foot-pounds of energy (sect. 225). This he must get from the food he eats.

88. Method of computing kinetic energy. The kinetic energy of a body is equal to the net amount of work that must be done on it to give it this energy. By taking a simple case it is possible to prove that this work is equal to

$$\tfrac{1}{2} \times (\text{mass of body}) \times (\text{square of its speed}).$$

Suppose that in throwing a baseball a constant force F acts while the ball is displaced a distance s. This force will produce a uniform acceleration a, the force and the acceleration being related by equation (21),

$$F = ma.$$

The work done by the force is, by equation (39),

$$W = Fs,$$

where s is the displacement while the force is acting. If the value of F from equation (21) is substituted in this equation,

$$W = mas.$$

But when a body moves with a constant acceleration (equation (12)),

$$s = \tfrac{1}{2}\,at^2.$$

Substituting this value of s,

$$\begin{aligned} W &= ma(\tfrac{1}{2}\,at^2) \\ &= \tfrac{1}{2}\,ma^2t^2. \end{aligned}$$

But the velocity the ball acquires when it has a uniform acceleration a is given by equation (11),

$$v = at.$$

If this is now substituted in the preceding equation, we have, as the net work done on the ball,

$$W = \tfrac{1}{2}\,mv^2. \tag{44}$$

In the C.G.S. system when mass is measured in grams, and velocity in centimeters per second, the energy as computed by this formula will be in *ergs*. For example, a bullet having a mass of 15 grams and a velocity of 700 meters per second will have a kinetic energy equal to

$$\begin{aligned} \tfrac{1}{2}(15) \times (70{,}000)^2 &= 3.675 \times 10^{10} \text{ ergs} \\ &= 3675 \text{ joules.} \end{aligned}$$

When B.E. units are used, the weight will probably be given in pounds. In that case it will be more convenient to change the form of equation (44), substituting *weight/32* for the mass m (see section 66). We then have

$$\begin{aligned} \text{Kinetic energy in foot-pounds} &= \tfrac{1}{2}\frac{\text{weight}}{32} \times v^2 \\ &= \frac{\text{pounds}}{64} \times v^2, \end{aligned} \tag{45}$$

where the velocity v is in feet per second. For example, a certain train has a weight of 1000 tons and a speed of 30 miles per hour. Hence

$$\text{Kinetic energy} = \frac{1000 \times 2000}{64} \times 44^2$$

$$= 60.5 \times 10^6 \text{ ft.-lb.}$$

$$= 30{,}250 \text{ foot-tons.}$$

89. Method of computing potential energy. The potential energy of anything is always measured by the net work done in giving it that energy or by the work it does in losing that energy. For example, in the case of a weight raised a certain height the potential energy is equal to the product of the weight in pounds times the height in feet. This is true because the product of weight times the height is the work done.

For B.E. units the potential energy of an elevated weight is

$$\text{Potential energy} = \text{weight} \times \text{height}, \tag{46}$$

giving foot-pounds of energy if the weight is in pounds and the height in feet.

If C.G.S. units are used, and the mass m in grams is given, the weight of the mass is mg dynes (sect. 66), and

$$\text{Potential energy} = mgh, \tag{47}$$

$$= 980\,mh \text{ ergs (approximately)}$$

when the height h is in centimeters.

In those cases where it is not a case of elevating a weight, the potential energy is determined either by computing the net work required to put the body in the condition where it has the energy or by the total work it does in losing that energy. A wound clock spring may have its energy determined by measuring the forces and distances through which these forces act when it is wound. The potential energy of coal is determined by experimentally measuring how much heat energy it develops when burned.

In general, the potential energy of a body can be determined either by the net work done in giving it that energy or by the total work it does when it loses that energy.

90. Power. When work is done we are interested not only in the total amount but also in the *rate* at which it is done. A ton of brick is hoisted 30 feet by an engine in 1 minute, and, in another case, by a man in a number of hours. The total effective work done is the same in the two cases, but the rate is different. We say that the engine, working at the greater rate, exerts more *power.*

Power is defined as the *rate of doing work.* If it takes a time t to do an amount of work W, then

$$\text{Average power} = \frac{W}{t}. \tag{48}$$

In the case of the engine's hoisting the brick, the power it exerted can be computed as follows:

$$\text{Work done} = 2000 \times 30 = 60{,}000 \text{ ft.-lb.}$$

Since this work is done in 60 seconds, the power is given by the equation

$$\text{Power} = \tfrac{60000}{60} = 1000 \text{ ft.-lb./sec.}$$

In this example the power is measured in *foot-pounds per second.* In a similar manner power can be measured in *ergs per second* or in *joules per second. Horse power* is a unit for measuring power, and its value is given by the equation

$$1 \text{ horse power} = 550 \text{ ft.-lb./sec.}$$

In the example given above, the answer was 1000 foot-pounds per second. When expressed in horse power this is

$$\tfrac{1000}{550} = 1.8 \text{ horse power.}$$

The beginner frequently fails at first to see that the horse power is not a unit of work; that is, it is *not equal* to 550 foot-pounds. It is equal to 550 foot-pounds per second, or 33,000 foot-pounds per minute. The horse power is an arbitrary unit, commonly used in all English-speaking countries. Whatever the origin of the name, it no longer means the power exerted by a horse: it is always equal to 550 foot-pounds per second.

Another unit of power is the *watt*. This well-known unit is used all over the world in connection with electrical apparatus.

$$
\begin{aligned}
1 \text{ watt} &= 1 \text{ joule per second}\\
&= 10{,}000{,}000 \text{ ergs per second}\\
1 \text{ kilowatt} &= 1000 \text{ watts}\\
1 \text{ horse power} &= 746 \text{ watts}^{*}
\end{aligned}
$$

Engineers often measure the power of engines and turbines in horse power, but for all electrical apparatus—lamps, dynamos, motors, etc.—watts and kilowatts are commonly used.

An equation which is convenient for those cases where one is concerned with velocity and force is the following:

$$\text{Power} = Fv, \tag{49}$$

where v is the velocity in the direction of the force F. An example of the use of this equation is easily found. The power exerted in pushing an automobile at a uniform speed is the product of the propelling force F and the speed v. If it takes 20 horse power to drive the car at a speed of 30 miles per hour, it is not difficult to compute the actual force pushing the car. Twenty horse power is equal to 11,000 foot-pounds per second, and 30 miles per hour is 44 feet per second. Since

$$\text{Power} = Fv,$$

$$11{,}000 = F \times 44.$$

Hence

$$F = 250 \text{ lb.}$$

If the car is running on a level road at a uniform speed, this force is numerically equal to the sum of the frictional forces. A relatively large force may be used in the case of trucks which run at a low speed. This force is the push the ground gives the rear wheels. Since the wheels push with an equal force on the ground, one can see why such cars are so destructive of roadways.

Equation (49) is easily proved. Power is equal to the work done in 1 second, and this is equal to the product of force and the distance moved in 1 second. Since v is the distance moved in 1 second, equation (49) follows at once.

---

* We should never say "horse power per second" or "watts per second."

In many instances it is more convenient to measure power than the total work. In such cases the power is first measured, and the work is found by the relation

$$\text{Work} = \text{average power} \times \text{time}. \tag{50}$$

Since the hour is often a more convenient unit of time than the second, work is frequently measured by the equation

$$\text{Work} = \text{horse power} \times \text{hours}. \tag{51}$$

The numerical result of this product is expressed in *horse-power hours*. Similarly, when the expression

$$\text{Work} = \text{kilowatts} \times \text{hours} \tag{52}$$

is used, the result is in *kilowatt-hours*, a unit of work or energy in common use.

91. Conservation of energy. A long time ago it was found that energy could not be *created* by any machine which man's ingenuity could invent. "Perpetual motion" machines are devices for the creation of energy. As long ago as 1775 the French Academy of Sciences decided not to waste its time listening to schemes for perpetual motion. Ingenious as many of these are, there is always something that the inventor has overlooked which prevents the machine from working. If all friction could be eliminated, there is no reason why a machine should not continue to run without any supply of energy after it was once started. But all machines which man has constructed have friction and are continuously losing energy; hence all our machines that run continuously must receive energy to compensate for this loss. If these machines do useful work, then still more energy must be supplied. Coal must be furnished to steam plants, moving masses of water to hydraulic turbines, and electrical energy to motors. Furthermore, it is a matter of experience that the more work a machine does, the more energy must be supplied to it. Careful experiments show that no machine can do more work than the energy supplied to it.

Energy cannot be created, but is it ever destroyed? Is the energy that disappears on account of friction really lost? It is now known that this energy is not destroyed; that it goes off in some other form of energy, most of it in heat. All experimental evidence points to the principle: *Energy is never created and never destroyed; the sum total of energy in the universe remains constant.* This is called the principle of the *conservation of energy.*

While it has been known for a long time that energy could not be created by machines, yet the generality of the principle was not accepted until about 1850. It was only after heat had been shown to be a form of energy that it was believed that energy was never destroyed by friction. Today it is believed that this principle of the conservation of energy applies everywhere—to atomic and molecular changes, to electrical phenomena, to biological phenomena, to all the processes of nature.

The principle of the conservation of energy is one of the most important generalizations of modern science, and we shall use it over and over again. The student will find that its use simplifies many problems which otherwise would be very difficult. Yet there is no direct way of proving this principle. We know it to be true in almost countless instances, and assume it always to be true. The principle is thus obtained by inductive reasoning based on experimental evidence.

92. Availability of energy. Notwithstanding the fact that energy is never destroyed, there is continually a decrease in its availability. Energy lost in friction goes into heat, and is dissipated beyond recovery. One reason that it is expensive to stop a train is that the kinetic energy possessed by the train is changed through friction into heat and is lost. In starting the train more coal must be burned and more work done by the engine to give the train the same kinetic energy. Much of the energy we obtain from coal and gasoline is promptly lost beyond recovery. All the energy used in "joy riding" is lost forever. Not only was there a tremendous loss of material in the battles of the World War, but enormous amounts of energy (chemical and mechanical) are gone beyond recall. The world's store of available energy has been greatly reduced by this appalling conflict.

The plea today for a greater conservation of our natural resources is to a large extent a plea for saving our available energy.

93. Sources of energy. Practically all our available energy has come from the sun. For ages sunlight acting on vegetable matter has stored energy in it. Part of this is now buried in the coal and oil fields. Vast amounts of energy from sunlight are stored each year in vegetation. From this, directly or indirectly, all animal life gets its food. Energy from the sun evaporates water and raises it to higher elevations. As the water flows back to the sea some of its energy is available for our use. The great convection currents of the atmosphere—the winds—obtain most of their energy from the heat given us by the sun. Enormous amounts of energy are possessed by the winds. Little of it, however, has as yet been converted to useful purposes. The utilization of the energy of water and of the winds will become more and more important engineering problems in the years to come.

There is another source of energy. The energy in the tides is produced at the expense of the rotational energy of the earth. According to theory this loss results in a slowing up of the rotational speed of the earth, but it is so slight an effect that, so far, the loss has not been detected.

94. Comparison of the C.G.S. and B.E. systems of units. The following table gives a comparison of the C.G.S. and the B.E. systems of units. The names of the *fundamental* units are in italics. The defining equation is given for the derived units.

System	Length	Mass	Time	Force	Work	Power
C.G.S.	*centimeter*	*gram*	*second*	$F = ma$ dyne	$W = Fs$ erg (joule)	$P = \frac{W}{t}$ erg/sec. (watt)
B.E.	*foot*	$m = \frac{F}{a}$ $= \frac{w}{32}$ slug	*second*	*pound*	$W = Fs$ ft.-lb. (foot-ton)	$P = \frac{W}{t}$ ft.-lb./sec. (horse power)

As shown in this table, the great difference between the two systems is in the choice of the fundamental quantities. In the C.G.S. system the unit of *mass* is a fundamental unit, while that of *force* is a derived one; in the B.E. system the fundamental quantity is *force* and not *mass*. In one the *gram* is the unit of *mass*; in the other the *pound* is the unit of *force*.

95. Summary. *Work* is defined and measured by the product of force and the displacement in the direction of the force. If a force acts on a body, and the displacement has a direction different from that of the force, either (1) the force times a component of the distance or (2) a component of the force multiplied by the distance the body is moved is used for computing the work done. The component which is used is always parallel to the other factor.

The unit of work in the B. E. system is the *foot-pound*, and in the C. G. S. system it is the *erg*. Since the erg is a very small unit, the *joule*, which is equal to 10^7 ergs, is often used.

In order that a body may gain *energy*, work must be done on it. If the body does work, it must lose energy by an amount equal to the work which it does. The *energy* a body possesses is measured by the amount of work which could be done if that body were to lose all its energy. The energy which a body has may be *potential* or it may be *kinetic*. A body has kinetic energy when in motion, and the magnitude of the kinetic energy is always equal to $\frac{1}{2}mv^2$. The potential energy is measured either by the net work which is done in giving the body that energy or by the total work done when the body loses its potential energy. In the case of a weight lifted against gravitational forces the potential energy is equal to the product of the weight times the height it is lifted.

Other units used for work or energy are the *calorie*, the *foot-ton*, the *kilowatt-hour*, and the *horse-power hour*. A table giving the numerical relations of the different units of work is found in section 81.

Power is the rate at which work is done. A rate of work of 550 foot-pounds per second, or 33,000 foot-pounds per minute, is equal to 1 *horse power*. A rate of 1 joule per second is equal to 1 *watt*.

The relation between power, force, and velocity is given by equation (49),

$$\text{Power} = Fv.$$

Experimental evidence indicates that energy is never created nor destroyed,—that the total amount of energy in the universe remains constant. In any system of isolated bodies (isolated in the sense that no energy can go into or out of it) the total energy is a constant. This principle is called the *conservation of energy*. The energy can change from one form to another, but the total amount in the isolated system will remain constant.

That the energy which we might use is becoming gradually less and less available is a fact of great economic importance. Energy "wasted" in friction is lost beyond recovery. Most of the energy obtained from coal and gasoline is lost forever.

It is important that everyone should understand the available sources of energy and have some idea of their practicability and duration. This question is one of continually increasing importance.

PROBLEMS

1. How much work is done in raising 40 cu. ft. of water to a tank from a well if the average height the water is raised is 30 ft.?

2. A mass of 1500 gm. is raised 200 cm. Compute the work in ergs.

3. If electrical energy costing 10¢ per kilowatt-hour is supplied to run a machine, how much would 10,000,000 ft.-lb. cost?

4. A force of 500 dynes acts for a distance of 2 m. on a mass of 1 kg. (*a*) How much acceleration will be given to the mass? (*b*) How much work will be done?

5. A laborer weighing 150 lb. carries a 10-pound hod containing 40 lb. of bricks up a vertical height of 20 ft. 5 times each hour. How much total work does he perform against gravity in an 8-hour day?

6. (*a*) How much work is required to roll a 150-pound barrel up an inclined plane 12 ft. long with an elevation of 4 ft.? (*b*) What is the direction of the least force required to roll the barrel up? (*c*) What is the magnitude of this force?

7. 100 lb. falls 30 ft. If all the energy could be converted into useful work, how much could be done?

8. A force of 30 lb. acting at an angle of 45° with the horizon drags a body a horizontal distance of 20 ft. Compute the work done.

9. (*a*) How much work is done in excess of that used in overcoming friction in driving a 3000-pound motor car 600 ft. up a 5 per cent grade (one that rises 5 ft. in each 100 ft. of road)? (*b*) What is the needed force?

10. A gun shoots a shell weighing a ton with a muzzle velocity of 2250 ft./sec. Compute the energy in foot-pounds.

11. A 3200-pound automobile, starting from rest, coasts down a hill 300 ft. long and 30 ft. high. If 50 per cent of the energy is lost in friction, how much speed should the car have at the bottom of the hill?

12. A constant force of 225 lb. acts on an automobile weighing 3200 lb. (*a*) Neglecting friction, what speed will the automobile attain on a level road in 10 sec. after the start? (*b*) How much work will be done in the 10 sec.? (*c*) How much kinetic energy will the car have at the end of the 10 sec.?

13. A mass of 400 gm. has a kinetic energy of 0.2 joules. (*a*) What is its velocity? (*b*) Through what distance must a force of 5000 dynes move in order to give it this energy?

14. (*a*) What momentum will a force of 4000 dynes give in 5 sec. to a mass of 400 gm. which starts from rest? (*b*) How much kinetic energy will the mass now have? (*c*) How much work has been done by this force?

15. A rifle weighing 7 lb. shoots a bullet, the weight of which is 0.0124 lb. (caliber 0.250), with a muzzle speed of 3000 ft./sec. (*a*) What is the momentum given to the bullet? (*b*) What is the kinetic energy in foot-pounds at the muzzle? (*c*) If no external force acted on the rifle, how much momentum would be given to it? (*d*) how much kinetic energy?

16. Find the horse power transmitted by a belt passing over a pulley, 4 ft. in diameter, which makes 300 revolutions per minute, the net force exerted by the belt on the pulley being 50 lb.

17. Prove that 1 horse-power hour is equal to 1,980,000 ft.-lb.

18. 30 tons of coal per hour are lifted 20 ft. by an elevator. What horse power is required?

19. Power is supplied at the rate of 25 watts. How many joules are supplied in 1 min.?

20. It requires 10 H. P. to drive a certain car at 20 mi./hr. If the total retarding force is three times as great, how many horse power will be required to drive the car at 30 mi./hr.?

21. The retarding force on an automobile traveling at 30 mi./hr. is 200 lb. What horse power must be furnished to maintain that speed?

22. A man weighing 150 lb. runs upstairs, rising 20 ft. in 5 sec. Compute the horse power he develops.

23. An automobile requires 50 H. P. to drive it 60 mi./hr. What is the total retarding force?

24. A man skating at a rate of 100 yd. in 20 sec. pulls a sled with a force of 10 lb. Find the horse power exerted on the sled.

25. It requires 500 H. P. to draw a certain train at the speed of 30 mi./hr. What is the numerical value of the force?

26. How much more horse power must a 3000-pound motor car exert to travel at 30 mi./hr. up a grade that rises 5 ft. in each 100 ft. of road than to travel on a level road at the same rate?

27. A windmill raises half a cubic foot of water 20 ft. each minute. If half the energy is lost through friction, what horse power is developed by the windmill?

CHAPTER VIII

SIMPLE MACHINES

Simple machines, 96. The lever, 97. Mechanical advantage, 98. Types of levers, 99. The inclined plane, 100. Weighing-machines, 101. The hydraulic press, 102. The effect of friction, 103. The efficiency of machines, 104.

96. Simple machines. This chapter gives an account of the applications of the principles of work and energy to some extremely simple machines. While practical machines are usually complicated, they involve the same principles as the simple ones here described. If these principles are understood, it is not difficult to study out the details of the more practical types. It is more important to understand fully and clearly the general principles of machines than it is to understand all the intricate details of a few commercial machines. The student must not expect to learn machinery from this chapter, but he should learn some fundamental principles.

97. The lever. In a previous chapter (sect. 10) the simple lever was explained as an example of the law of moments of force. It was shown that when the force and the weight are in equilibrium,

$$\text{Force} \times \text{distance from axis} = \text{weight} \times \text{distance from axis},$$

or

$$Fb = wa$$

and, as in equation (1),

$$F = \frac{a}{b} w.$$

This is called the law of the lever. A different method will now be used to deduce this law. The principle to be used is called the principle of work. It is based on the conservation of energy. *The principle of work asserts that in the absence of friction the work* *supplied to any machine is equal to that done by the machine.* Fig. 73 shows a lever which has been moved so that a weight w has been lifted a distance d; then

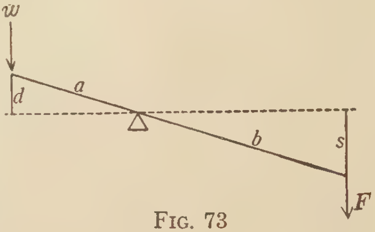


FIG. 73

$$\text{Work on } w = wd.$$

If s is the distance moved by the force F, then

$$\text{Work done by } F = Fs.$$

The principle of work states that these two are equal; that is,

$$Fs = wd.$$

Hence

$$F = \frac{d}{s} w.$$

The two triangles in Fig. 73 are similar; hence

$$\frac{d}{s} = \frac{a}{b}.$$

If this value of d/s is substituted in the preceding equation, equation (1) follows,

$$F = \frac{a}{b} w.$$

This is the law of the lever, which we have thus derived by a different method, the *principle of work.*

The force F, as given by equation (1), is just sufficient to balance w. It is not a sufficient force to start w; that is, to accelerate it. If w is to be given an acceleration, more force is required and more work is done; for then the load will be given kinetic energy.

The principle of work states that in frictionless machines

$$\text{Energy supplied} = \text{work done},$$

or $\quad \text{Applied force} \times \text{distance moved} = \text{load} \times \text{distance moved}.$

In using this principle, distances must always be measured parallel to the forces (see section 83).

98. Mechanical advantage. The ratio of the weight to the applied force, w/F, is called the *mechanical advantage.* It is a measure of the gain in force produced by the machine. While no energy can be gained by the use of a machine, yet there can be a gain in the magnitude of the force. In the case of the lever, if b is greater than a, the weight w is greater than the force F, and the mechanical advantage is greater than unity.

99. Types of levers. Either the principle of work or the law of moments of force (sect. 12) can be applied to any type of lever. In some cases the first is more convenient, while in other cases the second is better. This is illustrated by discussing several simple machines which may be regarded as types of levers.

The wheel and axle. In the wheel and axle a force F is applied tangentially to a wheel, balancing a larger force, w, tangential to the axle (Fig. 74). Let R be the radius of the wheel, and r that of the axle. By the law of moments (sect. 12),

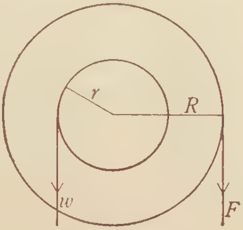


FIG. 74

$$FR = wr,$$

or

$$F = \frac{r}{R}w, \tag{53}$$

which is the law of the wheel and axle.

We may also derive this law by the principle of work. Suppose that the wheel has turned through 1 revolution; then the force F has moved a distance $2\pi R$, and the load w has been moved a distance of $2\pi r$. From the principle of work,

$$F \times 2\pi R = w \times 2\pi r,$$

or

$$F = \frac{r}{R}w.$$

The *mechanical advantage* of the wheel and axle is R/r.

There are many types of the wheel and axle in common use. In the steering-wheel of an automobile the axle transmits the force through gears. Belt-driven machines utilize in the pulleys and shafting the idea of the wheel and axle. One often sees shafting with pulleys of different sizes. Sometimes a large pulley mounted on a shaft is belted to the power supply, and a smaller pulley on the same shaft is connected to the machine to be driven. In another instance the source of the power is connected to the smaller pulley, and the machine to a larger one. The power or work is not changed by such a process, but the *mechanical advantage* is. The student should verify the following statement, which applies not only to shafting but to all types of machines: *When the mechanical advantage is greater than unity, the speed is reduced; when it is smaller than unity, the speed is increased.*

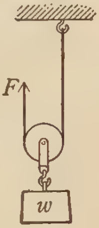


FIG. 75

The fixed pulley. A single fixed pulley is a form of lever with equal arms. There is no gain in force in the use of a fixed pulley, but it affords a convenient means of changing the direction of the force. Thus, in drawing a heavy weight up to a barn loft, it is more convenient to use a rope, pulleys, and a team than to lift the weight directly up.

The single movable pulley. In the case shown in Fig. 75 there is one pulley which moves when the weight is raised. If this pulley were raised 1 foot without moving the rope, there would be 2 feet of slack left in the rope. To take this up, the force F must move 2 feet. Hence, in using one movable pulley, the force F must move twice as far as the weight w. From the principle of work it evidently follows that

$$F = \tfrac{1}{2} w.$$

To use this method in any combination of pulleys it is necessary only to compute how far F moves when the weight w goes 1 foot. The distance that F moves when w moves 1 foot is numerically equal to the mechanical advantage. (Why?) (All friction and the weights of moving pulleys are neglected.)

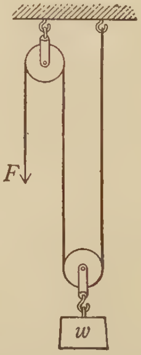


FIG. 76

Generally it is more convenient to use another method, especially when one continuous rope or cord is used. When this rope passes over a frictionless pulley the tension, or force, on the rope must be the same on the two sides of the pulley. Hence, neglecting friction and the weight of the rope, the tension must be the same at all points of a continuous rope. This fact can be used to derive the mechanical advantage of any combination of pulleys. For example, in Fig. 75 the weight w is supported by two ropes, so that the load on each is $\frac{1}{2}w$. Since the force F on the end of the rope must be equal to this tension,

$$F = \tfrac{1}{2}w.$$

The same result is obtained in the case of Fig. 76. In Fig. 77 the weight is supported by four ropes; hence the tension on each is $\frac{1}{4}w$ and

$$F = \tfrac{1}{4}w.$$

In Fig. 78 the load is supported by five ropes; hence

$$F = \tfrac{1}{5}w.$$

In commercial pulleys the friction is usually very large; hence the forces are found to be much larger than those given by the preceding formulas.

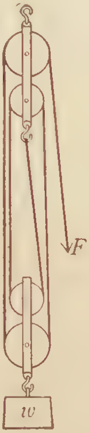


Fig. 77

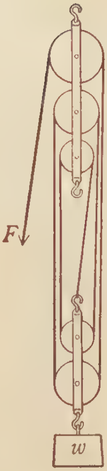


Fig. 78

The differential pulley. In the differential pulley a block is made of two pulleys, A and B, of slightly different diameters. These are fastened so that they always turn together. The grooves of the pulleys are so shaped that the chain, which is used in place of a rope, cannot slip. The endless chain passes over the pulley A, down under the pulley C, and then around B, with the slack hanging down (Fig. 79). The load w will be equally divided between the two chains which support it, so that the tension on each of these is $\frac{1}{2}w$. The one which passes over A will tend to turn A clockwise, exerting a moment $\frac{1}{2}w \times R$, where R is the radius of A; the one passing over B will tend to produce a counterclockwise moment, equal to $\frac{1}{2}w \times r$, where r is the radius of B. The *resultant* clockwise moment is

$$\tfrac{1}{2}w(R-r).$$

When r is very nearly equal to R, the resultant moment is very small. Usually the moment produced by the friction is greater than this, and the weight will remain stationary without a force F applied on the slack side of the chain. To lift the weight w, a

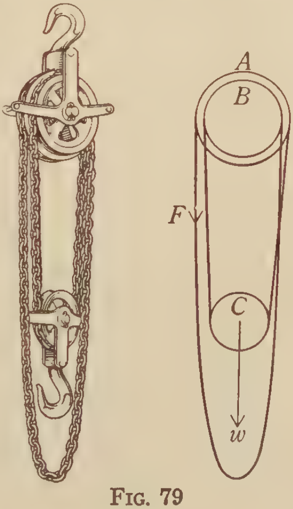


Fig. 79

force F must be applied, which will produce a counterclockwise moment, $F \times R$. Ignoring friction, for uniform speed,

$$FR = \tfrac{1}{2}\, w(R - r),$$

or

$$F = \frac{w}{2\,R}(R - r). \tag{54}$$

The mechanical advantage is

$$\frac{w}{F} = \frac{2\,R}{R - r}.$$

As r is very nearly equal to R, the mechanical advantage is very great. In practice the force required to lift w is at least twice as large as that given by equation (54), for the moment of force due to friction in the double pulley is at least as large as $\frac{1}{2}\, w(R - r)$, in order that the weight may not pull itself down. Equation (54) can also be derived by the principle of work by estimating the relative distances the force and load move.

100. The inclined plane. The law of the inclined plane will be derived by use of the *principle of work*.* The law to be proved is $F = \frac{d}{s}\, w$. In Fig. 80 the force is acting parallel to the length s of the plane. If the load is moved the entire length of the plane, the work applied is $F \times s$; but since the weight w is lifted a vertical height d, the work done is $w \times d$. Hence

$$Fs = wd,$$

or

$$F = \frac{d}{s}\, w. \tag{55}$$

Therefore the mechanical advantage is s/d.

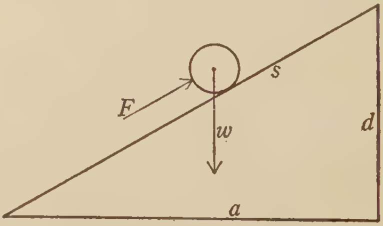


FIG. 80

It is left to the student to prove that for the case where the force is applied parallel to the base the mechanical advantage is a/d.

The *grade* of an incline, whether it be an ordinary inclined plane, or a railroad track, or a roadway, is defined as the ratio of the vertical height to the length of the incline. Usually the rise in 100 feet of incline is used. A 9 per cent grade, or 9 feet in 100, is quite steep for heavy teaming. For example, to draw a load of coal which, including the wagon, weighs 6000 pounds up a 9 per cent grade would require a force 540 pounds greater than that required to pull it on a level road. The student should verify this computation.

---

* See section 82 for another method.

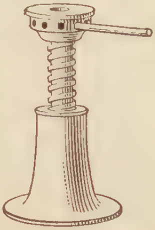


FIG. 81. The jackscrew

The various types of *screws*, such as jackscrews (Fig. 81), are examples of inclined planes. If a force F is applied at the end of a lever of a jackscrew, the energy supplied in 1 revolution is $F \times 2\pi r$, where r is the distance of the point of application of F from the center of the screw. If the distance between two adjacent threads of the screw, or the *pitch*, is d, then in 1 revolution the weight w would be lifted a distance d, and the work done would be equal to $w \times d$. Hence

$$2\pi r F = wd,$$

or

$$F = \frac{d}{2\pi r} w.$$

In practice the force is greater than this on account of the large amount of friction.

The propellers of boats or airplanes are special types of screws. The rotating blades push the water or air back, and the reaction drives the boat or airplane forward. The electric fan is a screw of the same type.

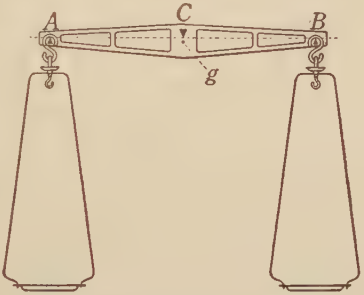


FIG. 82. The simple balance

101. Weighing-machines. We may regard a balance as a form of lever with equal arms, and readily conclude that the loads in the two pans must be equal when in equilibrium. However, we are usually interested in the sensibility of the balance; that is, how accurately we can weigh with it. While the complete theory is beyond the scope of this book, one or two points will be explained. The center of gravity of the beam, g (Fig. 82), must always be below the knife-edge C, which supports the beam. When the beam is tipped on account of unequal loading of the pans, the center of gravity is raised, moving from g to g' (Fig. 83). The work done in tipping the beam is equal to the weight of the beam times the height Dg. By the principle of work, this is equal to the excess weight in the pan times the vertical distance the pan moves. If the center of gravity is not near the knife-edge, tipping the beam through a small angle will raise the center of gravity a relatively large distance. In that case the work done is relatively large; hence, to tip the balance, a larger excess weight would be required in the pan. But if the center of gravity is very close to the knife-edge, it will be lifted a short distance, and the work done will be small. In that case only a small excess of weight will be needed. Therefore, in order to have the balance sensitive (that is, tipped by a small excess of weight), the center of gravity must be close to the knife-edge. Usually there is a small weight on the beam which can be moved up or down, so that the center of gravity can be raised or lowered and thus the sensibility be changed at will. In a somewhat similar way it can be shown that increasing the weight of the beam lowers the sensibility. It can be proved that the three knife-edges, A, C, and B in Fig. 82, should lie on the same straight line.

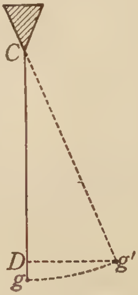


FIG. 83

In a familiar form of scales commonly seen in drug stores the two scalepans are placed above the beam. An interesting thing about this balance is that it does not make any difference where the load is placed on the scalepan. One would naturally infer, on first examination, that if equal weights were placed on the pans, with one weight at the extreme outer edge of the right-hand pan and the other on the inside edge of the left-hand pan, they would not be in equilibrium. The balance has two beams, AB and CD of Fig. 84, connected by knife-edges or pivots at A, B, C, and D to the rods which carry the scalepans. This gives a well-known geometrical device for parallel motion. The beams AB and CD always stay parallel, and the rods AC and BD are always parallel and vertical. Hence the scalepans remain horizontal. When a weight W is placed on the pan P, all points of P move equal distances downward. By the aid of the principle of work it can be seen that since the distance moved downward is always the same, the weight required to balance a given load will be the same for all positions on the pan. Owing to the fact that there is sometimes "lost motion" at

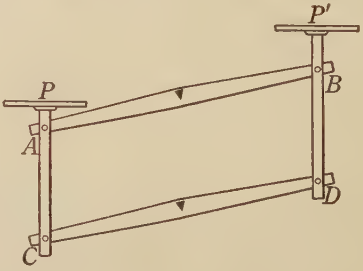


FIG. 84

the pivots A, B, C, and D, these conditions are not always exactly fulfilled. For this reason it is often possible to detect changes in the equilibrium if the weight is placed first on one side of the pan and then on the other.

102. The hydraulic press. The principle of the *hydraulic press* has been explained in section 27, but we can look at it from another point of view:

Let a be the area of the small piston (Fig. 85) and A that of the large one. If the smaller piston is pushed down a distance s, it will displace a volume of water sa, which will flow over into the other chamber and raise the larger piston a height d. The increase in volume of the water in this chamber will be dA. Since the two volumes are equal,

$$sa = dA,$$

or

$$\frac{d}{s} = \frac{a}{A}, \qquad (56)$$

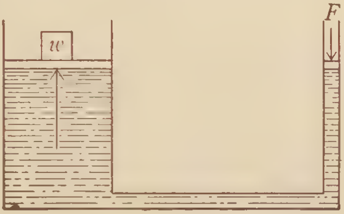


FIG. 85

which gives the ratio of the distances moved. We may now apply the principle of work. The work done by the force F is Fs, and the work done on the weight w is wd. As these must be equal,

$$Fs = wd,$$

or

$$F = \frac{d}{s} w.$$

By the aid of equation (56) this becomes

$$F = \frac{a}{A} w, \qquad (57)$$

and the *mechanical advantage* is $\frac{A}{a}$.

The same result can be obtained from the expression developed in section 27, where an entirely different line of reasoning is used. Agreements of this kind are common throughout the entire subject of physics. Such numerous cases of checking make us certain of the truth of the fundamental assumptions of mechanics.

103. The effect of friction. It should be clear to the student that while we have neglected friction in deducing the laws of different types of machines, in many cases friction must be taken into account in practice. In any practical case the following steps can be taken: first, the force is computed, neglecting friction; second, the effect of friction is estimated, and a correction applied. In the next chapter, methods of estimating the magnitude of frictional forces are discussed. In general, the effect of friction is to make the work done by a machine less than the energy or work supplied to it. Friction tends to lower what is called the *efficiency* of the machine.

104. The efficiency of machines. *The efficiency of a machine is the ratio of the work the machine does to the total energy supplied to it.* This term is used not only with the simple machines described here but with all classes of machines—engines, turbines, motors, dynamos, etc. The efficiency of a machine is commercially very important. When an engineer is deciding what type of machine to install, the efficiency, the initial cost, and the cost of maintenance are the most important items. It is often so necessary to know the numerical value of the efficiency that elaborate experiments are carried on to measure the work done and the energy supplied to the machine.

The term *efficiency* is used a great deal in a very loose sense; but there is nothing vague or indefinite about the term as applied to machinery, for it always has the definite meaning given in the definition above. The advantages of the method of precise measurement are obvious. When an attempt is made to improve the efficiency of a machine, there need be no guessing; the efficiency is measured, and the success or failure of the change is known. We hear a great deal about "efficiency methods" and "efficiency engineering." Often this is merely an attempt to apply the precise quantitative methods of mechanics to all sorts of problems.

PROBLEMS

1. A weight of 200 lb., 6 in. from the fulcrum of a lever, is balanced by a force 4 ft. from the fulcrum. Find (*a*) the applied force and (*b*) the mechanical advantage.

2. A bucket of water weighing 48 lb. is lifted by a wheel and axle. (*a*) If the axle is 6 in. in diameter and the wheel 3 ft., what force must be supplied to the rim of the wheel? (*b*) What is the mechanical advantage?

3. Two pulleys 24 in. and 8 in. in diameter are mounted on the same shaft. Power is supplied to the larger one, and a machine is belted to the smaller. What is the mechanical advantage of the shaft and pulleys?

4. A system of pulleys is so arranged that when the load is lifted 1 ft., the applied force moves 6 ft. (*a*) What is the mechanical advantage? (*b*) Draw a diagram of the arrangement of the pulleys.

5. How much force in dynes, parallel to the plane, would be required to support 10 kg. on a smooth plane 6 m. long and 1.5 m. high?

6. A wagon and contents weighing 4000 lb. is pulled up a hill 1000 ft. long and 50 ft. high. (*a*) How much extra work was done on account of the hill? (*b*) What was the extra force needed?

7. An inclined plane is 4 ft. high and 10 ft. long. How much force must be exerted parallel to the plane to roll a 300-pound barrel up the plane with a slow, uniform motion?

8. Given a block and tackle, each with two pulleys, what is the largest load that a man can lift (neglecting friction) if he can apply a force of 100 lb.? Make a diagram showing the arrangement.

9. Why is it more difficult to turn a screw with a small-handled screwdriver? Is it length or diameter of the handle that is important?

10. A painter hoists himself to the top of a structure by a block and tackle having two wheels in the upper (fixed) block and one in the lower (movable) block. Man and equipment weigh 200 lb. (*a*) What is the tension in the rope? (*b*) in the support of the fixed block?

11. What difference, if any, would it make if the painter of Problem 10 were hauled up by another man, standing on the ground below?

12. If the forces required to overcome friction in driving a 3000-pound car at 30 mi./hr. amounted to 300 lb. applied parallel to the road, how much force would be required to take the car up a 4 per cent grade at that speed?

13. A screw operated by a lever arm 36 in. long (measured from the center of the screw) has four threads to an inch. Neglecting friction, how large a load can be raised by a force of 10 lb.?

14. Water is supplied to a piston, which has an area of 30 sq. in., at a pressure of 40 lb./sq. in. Draw a diagram of a system of pulleys that will enable a total load of 2 tons to be lifted by the water pressure.

CHAPTER IX

FRICTION

Starting friction, 105. Sliding friction, 106. The angle of repose, 107. Power expended on account of friction, 108. Rolling friction, 109. Fluid friction: viscosity, 110. Viscosity in solids, 111. Viscosity in another sense, 112.

105. Starting friction. It is a matter of common knowledge that it requires a force to drag a body along a surface even when there is no acceleration. If we should measure the required force, we should find that it requires (1) a fairly definite force to *start* any particular body moving and (2) a smaller force to *keep* the body moving at a constant speed.

The retarding forces which in one case tend to prevent the starting and in the second case tend to stop the motion are called *forces of friction*, or often merely *friction.* These retarding forces can be determined by measuring the forces which must be applied to the body to overcome friction; thus the force required to keep a railroad train moving at a uniform speed on a straight level track is equal to the retarding forces of friction.

The force which is just sufficient to start a body moving is equal to what is called *starting friction.* The following facts regarding starting friction have been found by experiment:

1. It is different for different kinds of materials.
2. It is different for different conditions of the rubbing surfaces.
3. For any given body moving on a given surface it is approximately proportional to the force holding the body against the surface.

In Fig. 86, F is the force which is just sufficient to start a body moving, and F' is the force pressing the body against the surface. Usually F' will be the weight of the body. The third fact of the last paragraph can be expressed by the relation

$$F = \kappa F'. \tag{58}$$

The factor κ is called the *coefficient of starting friction.* The force required to start a body can be obtained by multiplying the bearing load by this factor. For example, if the coefficient of starting friction for a smooth iron block sliding on a horizontal polished iron surface is 0.3, the force required to start it when the block weighs 20 pounds is 6 pounds. If the block is loaded up to a total weight of 30 pounds, the starting friction becomes 9 pounds.

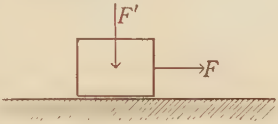


FIG. 86

If the body is resting on an inclined surface, the force required to start it sliding up will be a component of the weight of the body (sect. 82) *plus* the force F of equation (58). The force F is in that case the *extra* force which must be added on account of the friction. In the case of the incline the force F' is not the weight of the body but the component which is perpendicular to the surface of the incline.

Tables giving the coefficient of starting friction for various materials are published. From these tables one can find the coefficient of the particular materials one may be interested in, and, knowing the load, can make an estimate, by the aid of equation (58), of the starting friction.

106. Sliding friction. A smooth block of oak was drawn at a uniform speed over a horizontal polished surface of cast iron. By placing weights on the wooden block the force pressing the sliding surfaces together was changed. The force necessary to keep the block and its load moving uniformly was measured for different loads with the following results:

TOTAL LOAD F' IN POUNDS	FORCE F IN POUNDS	F/F'
4.84	1.14	0.24
7.04	1.58	0.22
9.24	2.10	0.23
11.40	2.60	0.23
13.60	3.10	0.23

It can be seen from these results that the force required to overcome the friction was approximately proportional to the load; or

$$F = fF'. \tag{59}$$

This type of friction is called *sliding* friction, and the constant f is called the *coefficient of sliding friction.* F' is the force pressing the sliding body against the surface. It is always in a direction perpendicular to the surface; when the surface is horizontal, as in the experiment just cited, and only in such a case, F' is equal to the weight of the sliding body. F is always the force necessary to overcome the friction; however, if the body is being accelerated or being raised up an inclined surface, F is only part of the applied force.

Equation (59) is only approximately true. For example, nothing is stated regarding the velocity of the sliding body. The value of the coefficient is not entirely independent of the velocity; for very small velocities and also for very large ones the sliding friction and, therefore, the coefficient are found to be larger. However, over quite a range of speeds the friction is practically independent of the speed.

It is also found by experiment that friction is very nearly independent of the area of the rubbing surface. This may be shown by measuring the force required to draw a weighted polished block of wood over a smooth surface. It is found that the friction is approximately the same whether the block is resting on its edge or on its side.

COEFFICIENTS OF SLIDING FRICTION

Wood on wood (dry surface)	0.25–0.5
Metals on oak (dry)	0.2–0.6
Metals on elm (dry)	0.2–0.25
Metals on oak (soapy)	0.2
Metals on metals (dry)	0.15–0.2
Leather on oak (dry)	0.27–0.38
Leather on metals (dry)	0.56

The coefficient varies so much with the nature of the surface that great reliance must not be put on the use of such a table as the foregoing. However, it is very seldom that one needs more than approximate estimates in computing friction.

Starting friction is always greater than *sliding* friction. A locomotive pulls a train by applying between the driving-wheels and the rails a force which is less than the starting friction. If, however, the initial force applied causes the wheels to slide, the force that can be obtained to move the train will be very much less, for then only a force equal to the sliding friction is available. Hence, in starting the train, care must be taken that the driving-wheels do not slip. A similar thing is true of the stopping of a train. The force that stops the train is a frictional force between rails and wheels. If the brakes are set too tight and the wheels are locked, the wheels will slide on the rails, and the train cannot be stopped so quickly. In walking or running one uses starting friction. If the foot slips and one is able to use only sliding friction, walking becomes very difficult. When an automobile turns a corner, the starting friction, not sliding friction, furnishes the necessary central force to overcome the centrifugal effect. However, if the car begins to skid, there is often nothing the driver can do to stop it. The sliding friction is often not large enough to change the direction of the motion of the car. For similar reasons the slipping of a belt on a machine means less power transmitted.

In a car wheel or wagon wheel there is sliding friction at the axle. When the axle is small and well lubricated and the wheel large, the retardation due to this may be very small. The force which must be applied to overcome this friction is much less than the actual friction, for the principle of the lever as applied to the wheel and axle enters in. The computation of the necessary force is not difficult. Suppose the wagon travels a distance such that the wheel makes 1 revolution, a distance equal to $2\pi R$, where R is the radius of the wheel. If the frictional force is F, the *work* done against this friction in 1 revolution is $F \times 2\pi r$, where r is the radius of the axle. The force applied to the wagon to overcome this friction must do, then, $F \times 2\pi r$ units of work in moving the wagon a distance $2\pi R$. The force will be equal to the work it does divided by the distance it moves. But as the distance it moves is $2\pi R$, the force necessary to overcome the axle friction is

$$\frac{F \times 2\pi r}{2\pi R} = \frac{r}{R}F. \tag{60}$$

Hence the force is equal to the ratio of the radii of the axle and wheel times the sliding friction of the bearing. The force will be less for larger wheels, provided the axle is not also made larger,

107. The angle of repose. If the surface is tilted through an angle θ until the body is just on the point of slipping, the angle θ is called the *angle of repose* (Fig. 87). The component of the weight parallel to the plane (F in the figure) is then just equal to the starting friction, but opposite to it. The force F' pressing the two surfaces together is not equal to the weight of the body but is that component of the weight which is perpendicular to the surface of the plane.

From the definition of the tangent of an angle, it is not difficult to prove that in this case the coefficient of starting friction is equal to $\tan\theta$.

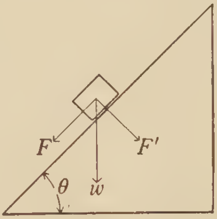


FIG. 87

108. Power expended on account of friction. Power is the work done per second; it is equal to the force times the distance moved per second. If F is the force necessary to overcome friction, and v the speed, or distance moved per second, then, by equation (49),

$$\text{Power} = Fv.$$

If the value of F from equation (59) be inserted in equation (49), then

$$\text{Power} = fF'v.$$

109. Rolling friction. When a steel wheel presses on a rail, the wheel is slightly flattened, and there is an indentation of the rail. Under these circumstances there are always forces which oppose the rolling motion. A wagon wheel sinks into a roadbed, even on a very hard road.

Rolling friction is usually much less than sliding friction. The great advantage of roller or ball bearings is due to the substitution of rolling friction for sliding friction. An inflated rubber tire should be softer for soft roads, for the flattening of the tire prevents its sinking into the road. Usually flattening increases the rolling friction, and for smooth, hard roads the tire should be hard.

Experiments show that in the case of wheels the retardation due to rolling friction varies inversely as the radius of the wheel: the larger the wheel the less the friction. The value of rolling friction is given by

$$F = \frac{f'W}{R}, \tag{61}$$

where F is the force applied at the center of the wheel to overcome the friction, W the load, R the radius of the wheel, and f' the value of the coefficient of rolling friction. The values used for f' will depend on the units in which R is measured. The following values of f' are for cases where the radius is measured in feet.

A car wheel on an iron or steel rail	0.0015–.002
A steel-tired wagon on a smooth, hard road	0.02
A steel-tired wagon on asphalt	0.012

In computing the force required to draw a wagon, in addition to the rolling friction we must include the sliding friction of the axle in its bearings. As an example take the case of a wagon with a total load of 4000 pounds and having all its wheels of the same radius, 1.5 feet, and with axles 3 inches in diameter. If we take the coefficient of sliding friction as .05, and the coefficient of rolling friction as .02, we have, for the total frictional force (see equations (60) and (61)),

$$\frac{r}{R}(.05\times 4000)+\frac{.02\times 4000}{R}$$
$$=\frac{1}{12}(.05\times 4000)+\frac{.02\times 4000}{1.5}=70\text{ lb.}$$

110. Fluid friction: viscosity. When water flows through a smooth, straight pipe there is friction; but it is not friction between the surface of the water and the surface of the pipe, for the water very close to the pipe does not move. The friction is entirely in the water. A similar thing is true when an object moves through water. In some liquids this friction may be quite large; for example, in thick oils and sirups. Gases also show fluid friction. A body moving through air not only carries along with it some air but also experiences a retarding force of friction. This internal friction of liquids and gases is called *viscosity*. Shot placed in such a liquid as sirup will fall slowly, owing to the great viscosity of the liquid. Particles of dust or small drops of water, such as in a fog or cloud, also fall slowly on account of the viscosity of the air. It is found that *the viscosity of fluids and gases increases with the speed* of the moving body. This gives a reason why a drop of water will fall through the air with uniform speed. At first, when it begins to fall, the weight of the drop will be greater than the friction; but as the speed of the drop increases, the friction will increase. The drop soon reaches that speed where the friction will be equal to its weight, and the resultant force becomes zero. It has been found possible to compute the size of small drops of water by measuring the rate at which they fall through the air and applying the complicated laws of viscosity.

It is worth while to explain why small drops fall at such low speeds; for example, the fine drops which constitute a cloud. The downward force acting on each drop is its weight, and the retarding force is the viscosity of the air. The weight is proportional to the volume of the drop, but the viscosity depends on the surface of the drop. Since the volume is proportional to the cube of the diameter and the surface to the square of the diameter, as the drop gets smaller its volume decreases much faster than its surface. Hence small drops have less weight as compared with the viscosity effect than do large ones. The result is that the retarding force on a very small drop becomes equal to its weight at low speeds. For the case where the diameter of a drop of water is one fiftieth of a millimeter, estimates show that it will fall at a uniform speed of 1.2 centimeters per second.

On the other hand, large raindrops must fall much faster in order for the retarding force to be equal to their weight. But there is a limit to the speed with which large drops of water can fall through the air. Whenever a drop exceeds a speed of about 8 meters (26 feet) per second, it is broken up into smaller drops, which cannot fall so fast.

Considerable attention has been given to the increase of fluid friction at high speeds. It is only at low speeds that the friction is proportional to the velocity. At higher speeds the friction varies approximately as the square of the speed, and at still higher speeds it increases more rapidly than that. As the speed of a train increases, the air friction increases very rapidly, until at high speeds it becomes greater than the other resistances, although these may also increase with the speed. It is this increase of friction with speed which makes it so expensive to run automobiles, trains, and ships at high speeds. It requires so much coal to drive a destroyer or a battle cruiser at its maximum speed that the vessel cannot carry enough to run at full speed for more than a comparatively few hours.

An increase in temperature lowers the viscosity of liquids but increases the viscosity of gases.

The laws of friction as applied to lubricated surfaces are also complicated. It appears, however, that when the surfaces are well lubricated, the friction is proportional to the relative velocities of the two surfaces.

111. Viscosity in solids. There is an internal friction in solids. The vibrations of a tuning-fork die out at a rate greater than can be explained by the rate at which energy is given to the air. A vibrating spring, or any vibrating solid, has internal friction which helps bring it to rest. The internal friction, or *viscosity*, in metals not only varies with different metals, but in the same metal it varies with the annealing. There is also internal friction in belts and cables used in driving machines.

112. Viscosity in another sense. The term *viscous* is often applied to solids which cannot resist the continued application of a force, although able to resist it for a short time. A piece of pitch struck by a hammer may fly into small pieces; yet if left alone long enough, under the continued action of its own weight, it will flow like a liquid. A piece of sealing-wax supported in a horizontal position by one end only will slowly bend under its own weight.

However, most metals show no signs of this sort of viscosity. The sharp outlines of the impressions on very old coins, and the preservation of old pieces of metallic art, show that this type of viscosity does not exist to any appreciable amount in these metals.

PROBLEMS

1. It required a horizontal force of 50 lb. to drag a 200-pound weight over a smooth floor at a uniform speed. Compute the coefficient of friction.

2. A force of 150 lb. is required to pull with uniform speed a sled which with its load weighs 2000 lb. Find the coefficient of friction.

3. A horizontal force of 40 lb. gives to a 320-pound weight a horizontal acceleration of 2 ft./sec.^{2} (*a*) What is the frictional force? (*b*) What is the coefficient of friction?

4. If the coefficient of friction is 0.2 and the load is 1000 lb., how much work will be done in moving the load at constant speed for 200 ft.?

5. If the coefficient of starting friction between the driving-wheel of a locomotive and a rail is 0.3, what minimum weight must the driving-wheel carry in order that the pull exerted through this wheel on the train may be 1 ton?

6. A 3000-pound automobile will coast down a 5 per cent grade at a constant speed of 20 mi./hr. What is the retarding force on the car?

7. On a horizontal road a force of 200 lb. gives a 3200-pound car an acceleration of 1 ft./sec.^{2} What is the retarding force?

8. A mass which weighs 400 lb., starting from rest at the top of an incline 30 ft. high, acquires a speed of 20 ft./sec. at the bottom. How much energy was lost on account of friction?

9. A force of 200 lb. was necessary to slide a 300-pound box up an incline 12 ft. long and 4 ft. high. Find the force of the friction.

10. A shaft 1 in. in diameter has a total load of 250 lb. pushing it against its bearings. Compute the force necessary to overcome the friction if the coefficient of friction is 0.05 and if the force is applied to the rim of a pulley 18 in. in diameter, which is mounted on the shaft.

11. A wagon with a total load of 5000 lb. has wheels 3 ft. in diameter and axles 2 in. in diameter. If the coefficient of sliding friction at the axle is 0.15 and the coefficient of rolling friction on a certain hard road is 0.025, what will be the necessary horizontal force, applied at the axle, to keep the wagon moving at a constant speed?

12. What will be the force required to keep a 100,000-pound car moving on a level track if the coefficient of sliding friction at the axle is 0.05 and the coefficient of rolling friction is 0.002, and if the diameters of the wheel and axle are 30 in. and 2 in. respectively?

13. (*a*) Find the work done by a horse in hauling a sled weighing 1600 lb. for a mile along a level road if the coefficient of friction is 0.05. (*b*) If the horse travels 3 mi./hr., what power does he deliver?

14. The frictional force on the surface of a 1-inch shaft is 5 lb. How much power is expended in overcoming this friction if the shaft makes 300 revolutions per minute?

CHAPTER X

VIBRATORY MOTION

Vibratory motion, 113. Simple harmonic motion, 114. Another definition of simple harmonic motion, 115. Proof of the equality of the two definitions, 116. Period, amplitude, phase, 117. The fundamental equation of simple harmonic motion, 118. The simple pendulum, 119. The physical, or compound, pendulum, 120. The torsion pendulum; angular simple harmonic motion, 121. Vibrations in musical instruments, 122.

113. Vibratory motion. Ever since Galileo watched a swaying chandelier in a cathedral in Pisa and, partly by observation and partly by guesswork, found one of the laws of its motion, the vibrations of pendulums have been of practical and theoretical interest. The oscillations of a pendulum are only one case of what is called vibratory motion. In all musical instruments sound is produced by something which has a vibratory motion. There is some evidence to show that Pythagoras (500 or 600 B.C.) worked out some of the simpler laws of the vibrations of the strings of musical instruments.

The principles which we shall study apply not only to cases such as have been mentioned but also to wave-motions; for wave-motions, such as waves traveling over the surface of water, are a special type of vibratory motion. They apply also to light-waves and electrical waves, where the vibrations are not those of a material medium. The detailed study of alternating electric currents, the kind ordinarily used in lighting-circuits, is to a large extent an application of the principles of vibratory motion.

That type of vibratory motion which, all things considered, is the simplest is called *simple harmonic motion.* The use of the term *harmonic* arose on account of the fact that the first detailed study of this subject was made in connection with the study of musical vibrations.

114. Simple harmonic motion. If a mass which is suspended by a spiral spring is pulled down, and then released, it will vibrate up and down with simple harmonic motion. In this case it is not difficult to state the facts relating to the magnitude and direction of the resultant force acting on the mass.

1. When the mass is pulled down, there is an upward force acting; if the mass is pushed up, there is a downward force. There is always a "return" force tending to restore the mass to its initial position of rest, the position of equilibrium.

2. It is well known that when a spring is stretched or compressed, the reacting force is proportional to the change of length of the spring. It follows from this that the return force produced by the spring is directly proportional to the distance the mass is displaced from its position of equilibrium.

These facts are characteristic of simple harmonic motion. Hence we may define simple harmonic motion as follows: *In simple harmonic motion there is at each instant a restoring force proportional to the distance the body is displaced from its equilibrium position.*

Whenever it is known that the restoring force is proportional to the displacement, then it is known that the vibrations are simple harmonic vibrations. The laws of elasticity tell us that this is true in the case of the spiral spring, in the case of thin steel rods clamped at one end, and in many other cases involving elastic forces. In the case of a pendulum, or a boy in a swing, it is proved in section 119 that the restoring force is proportional to the displacement, provided the amplitude of motion is small. Hence in these cases the motion is simple harmonic motion.

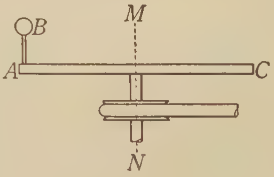


Fig. 88

115. Another definition of simple harmonic motion. A ball B is mounted on any suitable support AC, which can be rotated about an axis MN (Fig. 88). The type of rotator commonly used on the classroom table is satisfactory. When the ball is revolving around the axis MN, an eye situated in the plane in which the ball is moving will see only a vibratory motion. Or the rotating ball may be placed in a beam of light in such a position that its shadow will have only vibratory motion. The shadow S (Fig. 89) of the ball B will move in a straight line between the limits P_1 and P_2. If the ball is moving with a *constant angular speed*, the shadow will have simple harmonic motion.

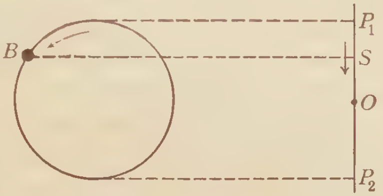


Fig. 89

Simple harmonic motion is the motion of the projection on a straight line of a point which moves with uniform speed in a circle, the projected motion being in the plane of the circle.

In the illustrations of simple harmonic motion given in the last section there was no circular motion. Usually one must imagine this circular motion. In Fig. 90 the vibrating mass m moves up and down from A to E. One can imagine a point rotating around the circle at such a speed that it always stays in line with the weight. When m is at B, the point is at B'; when m is at C, the point is at C'; and so on. From the definition just given, if the mass m does have simple harmonic motion up and down, and if the point is moved around the circle with such a speed that it will always be opposite the mass m, then the point *must move around the circle with uniform speed.*

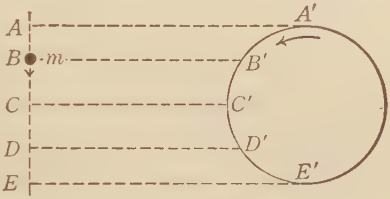


Fig. 90

We have had two different definitions for simple harmonic motion. We shall next prove that these two definitions are equivalent.

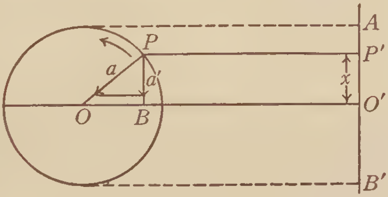


Fig. 91

116. Proof of the equality of the two definitions. We shall start with the second of the definitions given above and deduce the other. When the point P (Fig. 91) rotates around the circle with constant speed, its projection P' moves up and down the line AB' with simple harmonic motion. In section 77 it was shown that a body moving in a circle with uniform speed has a constant acceleration toward the center equal to v^2/r (equation (36)), where v is the constant speed around the circle, and r its radius. Since the projection P' has the same vertical velocity as P, it has the same vertical acceleration as P. If a is the acceleration of P, then a', the vertical component, is the acceleration of P'.

From the similar triangles in the figure,

$$\frac{a'}{a}=\frac{PB}{PO}=\frac{x}{r},$$

where x is the distance of P and of P' from the line OBO', and r the radius of the circle. The acceleration of P' is then

$$a'=a\frac{x}{r}.$$

But, from equation (36), $$a=\frac{v^2}{r}.$$

Hence $$a'=\frac{v^2}{r^2}x. \tag{62}$$

As shown in the figure, x is the *displacement* of P'. Since both v and r are constant, this equation shows that the acceleration a' of P' is proportional to the displacement x. If a mass m be placed at P' and be given the same motion as P', then it too must have an acceleration of a'. If we multiply both sides of equation (62) by m, we have

$$ma'=m\frac{v^2}{r^2}x.$$

But mass times the acceleration is equal to the force which must be applied to the mass to produce the acceleration. Hence

$$\text{Force}=m\frac{v^2}{r^2}x. \tag{63}$$

This is the force acting toward O', the so-called restoring force. On the right-hand side of this equation all the quantities are constant except the displacement x. Hence *the force is proportional to the displacement.* We have, then, started with one definition and derived the other.

117. Period, amplitude, phase. The *period* of a vibratory motion is defined as the time required for a round trip,—the time of a complete vibration.

The *amplitude* of a vibratory motion is the maximum displacement of the vibrating body from the center. In Fig. 91, AO' is the amplitude.

Phase is a term used in stating where the body is at any instant in its vibration, and what its direction of motion is. Two pendulums may have the same period and the same amplitude, and yet they may not be swinging together. One may be going to the left when the other is going to the right. If so, they have not the same *phase.* Phase is usually measured by reference to a circle; for example, the phase of P' (Fig. 91) in its vibratory motion can be measured by the angle POB.

118. The fundamental equation of simple harmonic motion. The period of P' (Fig. 91) is the time required for P to go once around the circle. Since the distance around is $2\pi r$ and the speed is v, the period is given by the equation

$$T = \frac{2\pi r}{v}. \tag{64}$$

If we substitute in equation (63) the value of v/r obtained from equation (64), we have, as the value of the restoring force,

$$F = \frac{4\pi^2 m}{T^2} x. \tag{65}$$

This equation applies to all cases of simple harmonic motion (S.H.M.) and may be called the *fundamental equation* of that motion. The meaning of each quantity in this equation should be clearly understood. F is the restoring force which is always acting toward the center or position of equilibrium; m is the mass of the vibrating body; x is the instantaneous distance from the center, or the displacement at the instant the return force is F; and T is the period, or time of one complete vibration.

119. The simple pendulum. By a *simple* pendulum is meant a small, heavy mass suspended by a fine cord or wire. This is called a simple case because the weight and mass of the suspending wire and the frictional force of the air can be neglected. If the volume of the suspended body is small, we can also assume that its mass is all located at one place, the center of gravity. To avoid other difficulties, we shall also take the case of a pendulum swinging through a small amplitude. For such simplified conditions it will be proved that

$$T=2\pi\sqrt{\frac{l}{g}}, \qquad (66)$$

where T is the *period*, or time of one complete vibration, l the length of the pendulum, and g the acceleration of a freely falling body. From this formula or law one can conclude that

1. The period varies directly as the square root of l. To make the period twice as great the pendulum must be four times as long. An increase of temperature makes l longer and thus makes the period greater. Hence an increase of temperature tends to make a pendulum clock run more slowly.

2. The period varies inversely as the square root of g. On a mountain, where g is less, the period is greater.

3. The period is independent of amplitude, provided the amplitude is small. If the angle through which the pendulum swings is not small, for accurate work one can use the formula

$$T=2\pi\sqrt{\frac{l}{g}}\cdot\left(1+\frac{1}{4}\sin^2\frac{\theta}{2}+\frac{9}{64}\sin^4\frac{\theta}{2}\right),$$

where θ is the *angular* amplitude. When the angular amplitude is 5 degrees, the error in equation (66) is about one fiftieth of 1 per cent. The fact that the period of a pendulum is nearly independent of the amplitude makes it less important for the pendulum of a clock to have a constant amplitude, but high-grade clocks are so constructed that the amplitude of the pendulum will remain nearly constant.

In a physics laboratory the simple pendulum affords a method of determining g, the acceleration of freely falling bodies. If the quantities T and l, appearing in equation (66), are measured, g can be computed.

The method which will be used for deriving the formula for a simple pendulum is first to prove that we have a case of simple harmonic motion. When this is done, the fundamental equation of simple harmonic motion, equation (65), is used. In Fig. 92 the weight w of the pendulum is represented by the vector CH, where C is the center of gravity of the small ball. In the figure this vector is resolved into the two components CE and CD. The force CE is parallel to the supporting cord and does not have any effect in moving the pendulum. The force CD is perpendicular to the string and is the effective force acting on the pendulum. If the displacement BC is very small compared with the length of the pendulum, we can assume that the pendulum ball moves along a straight line BC rather than along an arc. The force CD, or F, is the force acting toward the center. In the figure the triangles ACB and CEH are similar; therefore

$$\frac{EH}{CH}=\frac{CB}{AB}.$$

But EH is equal to CD, or F, the restoring force; CH is equal to the weight of the mass m and is therefore equal to mg; CB is equal to the displacement x; and AB is equal to the length l of the pendulum.

Hence
$$\frac{F}{mg}=\frac{x}{l},$$

or
$$F=\frac{mg}{l}x. \qquad (67)$$

On the right-hand side of this equation all the quantities are constant except x. Hence the restoring force is proportional to the displacement; and since this is a definition of simple harmonic motion, we must have simple harmonic motion here. (If the pendulum were vibrating through a large angle, this proof would not be sound. Why?) Another equation for F is the fundamental equation of simple harmonic motion given by equation (65).

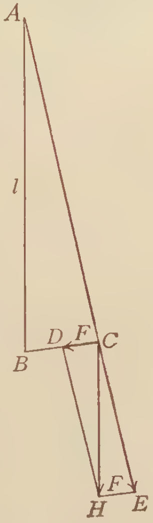


FIG. 92

Equating the right-hand sides of equations (67) and (65), we have

$$\frac{mg}{l}x=\frac{4\pi^2 m}{T^2}x.$$

Solving for T^2 and canceling common factors,

$$T^2=4\pi^2\frac{l}{g}.$$

Hence (equation (66))
$$T=2\pi\sqrt{\frac{l}{g}}.$$

120. The physical, or compound, pendulum. In the theory of the simple pendulum it is assumed that the weight and mass of the cord or wire may be neglected and that the ball is so small that its mass may be regarded as concentrated at its center of gravity. Most pendulums depart greatly from these simple conditions. For example, a long cylindrical bar is often used in our laboratories. One which does not satisfy these simple conditions is called a *physical*, or *compound*, pendulum.

It can be proved that a physical pendulum has the same period as a simple pendulum of length l if

$$l = \frac{I}{mR},$$

where m is the mass of the physical pendulum, R the distance of its center of gravity from the point of support, and I the *moment of inertia* of the pendulum (see section 132). If this value of l is substituted in equation (66), for the period of the physical pendulum we obtain the equation

$$T = 2\pi\sqrt{\frac{I}{mRg}}. \qquad (68)$$

A more accurate determination of the value of g can be made with a physical pendulum than with a simple one. The method, however, demands some means of determining I.

121. The torsion pendulum; angular simple harmonic motion. When a rod which is supported by an elastic wire (Fig. 93, *a*) is moved in a horizontal plane so as to twist the wire, and is then released, it will oscillate back and forth through an angle, marked θ in Fig. 93, *b*. In Fig. 93, *b*, the wire is perpendicular to the plane of the paper, and is fastened to the rod at the point O. The motion of the rod is simple harmonic motion, the rod moving not in a straight line but through an angle. For this reason this type is called *angular simple harmonic motion*. A rod vibrating in this manner is an example of a *torsion pendulum*. The balance wheel of a watch is a torsion pendulum; but instead of using the torsion of a straight wire, it is the torsion of the hairspring which furnishes the restoring moment of force.

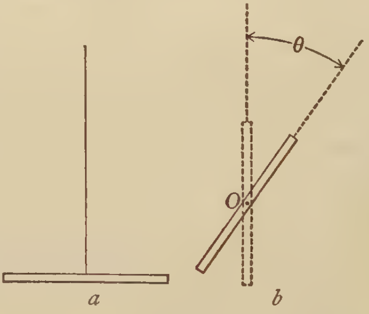


FIG. 93

122. Vibrations in musical instruments. In a stringed musical instrument the strings vibrate back and forth in harmonic motion. Usually a string gives out a certain tone together with a number of higher-pitched tones called overtones. But when a string gives out a pure tone, a tone of only one pitch, then the string vibrates as shown in Fig. 94, and each particle vibrates with simple harmonic motion.

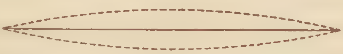


Fig. 94

In wind instruments there are air columns which are set in vibration. Usually the motion is quite complicated, and tones of more than one pitch are emitted. In this case and also with vibrating strings the vibrations can be regarded as due to the superposition of a large number of simple harmonic motions of different periods.

PROBLEMS

1. A body vibrating with simple harmonic motion has an amplitude of 10 cm. and makes 3 vibrations per second. Compute the velocity at the middle of its swing.

2. If the mass of the vibrating body of Problem 1 is 200 gm., find the restoring force (*a*) when it is displaced 2 cm. from its position of equilibrium, (*b*) when it is displaced 4 cm., and (*c*) when it is displaced 6 cm.

3. A weight of 5 lb. is hung from a spring. It is found that a force of 2 lb. is required to displace it 1 in. What period will it have when set in vibration?

4. A mass in simple harmonic motion makes 2 vibrations per second with an amplitude of 2 cm. Compute the acceleration (*a*) at the end of the vibration, (*b*) when the displacement is 1 cm., and (*c*) at the middle of the swing.

5. If the mass of Problem 4 were 300 gm., how much force would be required to hold it with a displacement of 1 cm.?

6. The end of a prong of a tuning-fork which makes 200 vibrations per second has an amplitude of 1 mm. Compute the maximum velocity of the end of the prong.

7. What will be the length of a pendulum that has a period of 1 sec. if g is equal to 980 cm./sec.^{2}?

8. A simple pendulum consisting of a small lead ball is hung by a fine thread. If the length of the pendulum is 1 m., and g is equal to 980 cm./sec.^{2}, what should be the period of the pendulum?

9. In a laboratory experiment the period of a simple pendulum, the length of which was 90 cm., was found to be 1.90 sec. Compute the value of g.

10. A pendulum clock which runs correctly (period, 1 sec.) at a place where g is equal to 980 cm./sec.^{2} is taken to Denver, where g is equal to 979.6 cm./sec.^{2} (*a*) What will be the period of the pendulum? (*b*) How many seconds will be gained or lost in 24 hr.?

11. Compute the kinetic energy of the mass in Problems 1 and 2 as it moves through the position of equilibrium.

12. A mass of 784 gm. is suspended by a light spring. An additional load of 20 gm. is required to displace this mass downward 1 cm. more. What will be the period when the 784-gram mass is vibrating? ($g = 980$ cm./sec.^{2})

CHAPTER XI

ROTARY MOTION*

Axis of rotation, 123. Angular speed, 124. Relation between radians per second and linear speed, 125. Relation between revolutions per minute or second and angular speed in radians per second, 126. Angular acceleration, 127. Gyroscopic motion, 128. Vector addition of angular velocities, 129. Resolution of angular velocities; components of rotation of the earth, 130. Effect of the rotation of the earth on the motion of projectiles and winds, 131. Kinetic energy of rotating bodies; moment of inertia, 132. Radius of gyration, 133. Formulas analogous to those of linear motion, 134. Conservation of angular momentum, 135. Total kinetic energy, 136.

Nearly everyone has watched a spinning top and wondered at the apparently mysterious force which kept it from falling. But this is only one instance; there are numerous cases of rotating bodies, many of practical importance, which require special methods to explain them. It is well worth while to learn the methods of analyzing problems relating to bodies in rotation.

123. Axis of rotation. In the case of a rotating body it is always possible to find a line which is not moving. This line is called the *axis of rotation.* In a flywheel turning on a shaft the axis is a line through the center of the shaft. Usually it is a simple matter to locate the axis of a rotating body.

124. Angular speed. The linear speed of a rotating body is an indefinite quantity; it is zero at a point on the axis, and increases the farther the point is away from the axis. It thus has all values between zero and the maximum value which points on the outside edge have. Hence it is not convenient to measure or state the speed of rotation of a body in terms of linear speed. Instead, the speed is measured in terms of the *angular speed.* For example, the average angular speed of the body in Fig. 95 is equal to the angle through which the line *AB*, which is drawn perpendicular to the axis, moves in 1 second. When angular speed is measured in *radians per second* (see section 77), it will be denoted by the Greek letter ω. Angular speed may be measured also by the number of *revolutions per second* (R.P.S.) or by the number of *revolutions per minute* (R.P.M.).

---

* This chapter may be omitted in a short course.

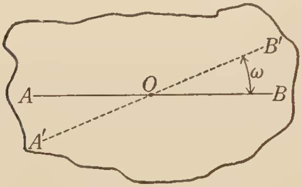


FIG. 95

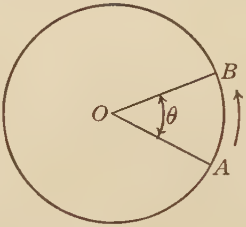


FIG. 96

125. Relation between radians per second and linear speed. It is often convenient to be able to reduce linear speed to angular speed. For example, one might know the linear speed of a belt, and hence that of the rim of the pulley, and wish to compute the angular speed of the pulley.

One formula may be deduced as follows: Take some point, *A*, in a body which rotates about an axis through *O* (Fig. 96). Let *B* be the position of this point after an interval of time τ. Then the arc *AB* is equal to $v\tau$, where v is the average linear speed. When the angle *AOB* is measured in radians (see equation (32), sect. 77),

$$\angle AOB = \frac{\text{arc } AB}{r} = \frac{v\tau}{r},$$

or

$$\theta = \frac{v\tau}{r}.$$

But

$$\theta = \omega\tau,$$

where ω is the angular speed in radians per second.

Hence

$$\omega\tau = \frac{v}{r}\tau,$$

or

$$\omega = \frac{v}{r}, \tag{69}$$

when ω is measured in radians per second.

126. Relation between revolutions per minute or second and angular speed in radians per second. The angular distance traveled by a body in 1 complete revolution is 360 degrees, or 2π radians. Hence if the body makes 100 *revolutions per second* (R.P.S.), the angular distance traveled in 1 second is $100 \times 2\pi$ *radians.* In general,

$$2\pi \times \text{R.P.S.} = \textit{angular speed in radians per second} = \omega$$

and
$$2\pi \times \frac{\text{R.P.M.}}{60} = \omega.$$

127. Angular acceleration. Angular acceleration is defined in a manner similar to linear acceleration. Angular acceleration is the rate of change of angular velocity. The change may be in the magnitude of the angular velocity, or it may be a change of the direction of the axis of rotation, or it may be both.

128. Gyroscopic motion. If one holds a rotating bicycle wheel by the axle and tries to turn around, or by any method tries to change the *direction* of the axle, he will feel a curious effect. As long as the wheel is spinning, the axle will not turn in the expected direction but will try to move in a direction approximately at right angles to the direction it would move if not spinning. For example, *WW′* of Fig. 97 represents a wheel spinning about an axis *AA*. If one should try to move the upper end of this axis out toward the eye, instead of moving in the expected direction this end of the axle will tend to move sidewise in a direction at right angles to the applied force. The same effect is shown in a spinning top. Instead of falling when it is leaning, its axis moves sidewise. An elementary explanation of this action will first be given in terms of linear velocities, and later by a method which uses rotational vectors.

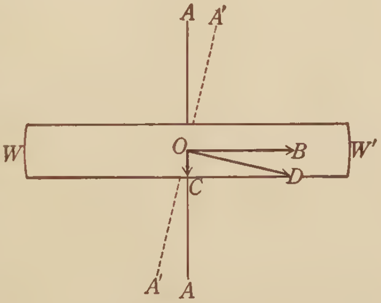


FIG. 97

In Fig. 97 the wheel WW' is rotating about the line AA as an axis. The linear velocity of a point O on the front side of the outside edge is represented by the vector OB. Imagine that the wheel is given an impulse which tends to move the upper end of the axis AA toward the eye. This impulse will give to the point O a linear velocity which will be downward (represented in Fig. 97 by the vector OC). The resultant motion of the point O is represented by the vector OD, which is the vector sum of OC and OB. In order that the point O may move in the direction OD, the axis of the wheel must shift to the position $A'A'$, which lies in the plane of the paper. In this case a force is applied which tends to *pull the upper end of the axis out toward your eye*; instead of moving in that direction, *it moves at right angles to it.* A more complete explanation would have to take into account other parts of the rotating body, but this will at least show the "reasonableness" of the observed result. After all, this is not an unusual occurrence. When a boy whirls a stone around on the end of a string, he pulls on the string; but this does not pull the stone toward him but keeps it moving in a circular path. The attractive force the earth exerts on the moon does not move the moon nearer the earth. When we apply a force, we must take into account the velocity the body already has before we can state what the resultant motion will be. In the case of the wheel, at first thought we do not take into account its velocity. After all the motions are considered, there is no longer any mystery about the behavior.

129. Vector addition of angular velocities. An angular velocity can be represented by a line or vector. The *length* of the line is always drawn proportional to the magnitude of the angular velocity, and the *direction* of the vector is the direction of the *axis* of rotation. The following simple rule will tell the direction in which the vector must be pointed. Grasp, in imagination, the axis of the rotating body with the right hand, the fingers pointing in the direction of rotation, with the thumb pointing along the axis. The direction the thumb points along the axis is the direction the vector should point. Rotational velocities *may be added and subtracted by the vector methods of Chapter I.* Although this is an important theorem, a proof of it will not be attempted here.

Vectors which have been drawn in accordance with the right-hand rule may be used to explain the motion of the bicycle wheel, a spinning top, or the interesting motions of any gyroscope. In using this method it is necessary to keep clearly in mind the direction of the axis of rotation and the direction of the axis about which one tries to rotate the spinning body. In Fig. 98 the vector *OA*, which is drawn parallel to the axis *AA* of Fig. 97 and in the direction given by the right-hand rule, represents the angular velocity around the axis *AA*. Let us suppose that an angular velocity about a horizontal axis in the plane of the paper is added to the wheel of Fig. 97. The vector *OM* is drawn in the proper direction to represent this added rotational velocity (the student should verify this direction). The resultant of these two vectors is *ON*. Since rotational vectors are always parallel to the axis of the rotation, this shows that the axis must move from a direction parallel to *OA* to one parallel to *ON*, or, in Fig. 97, from *AA* to *A′A′*. This method then gives the same result as before, but it has the advantage, to those familiar with it, of being a shorter one.

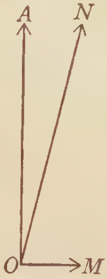


FIG. 98

This method of adding angular velocities and of finding the resulting direction of the axis of rotation may be applied to a spinning top leaning as shown in Fig. 99. In this case there is a torque due to the weight *w*, which may be taken as acting at the center of gravity. This torque tends to add an angular velocity about a horizontal axis *OB*, which is perpendicular to the line *OX*. The angular velocity of the top is represented in the figure by the vector *OA*. The angular velocity which would be added in a short interval of time by the weight is represented by a vector drawn in the direction *OB*. The sum of these two vectors is a line from

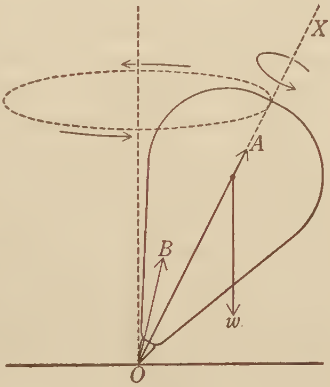


FIG. 99

the point O extending in an oblique direction away from the eye, behind the vector OA. This gives the new direction of the axis of the resultant spin. Hence the upper end of the axis of the top will move away from the eye. But as it moves, the direction of the axis OB, which gives the direction of the added vector, will change, for it must remain perpendicular to OX. The result is that the upper end of the axis OX will move in a circle, shown by the broken line in the figure. This motion of the axis of rotation of a body is called *precession*. The effect is quite analogous to the case of a small mass whirled on the end of a string around a circular path. In that case a central force is continually adding a velocity in a direction at right angles to the velocity of the mass, with the result that the velocity is continually changing in direction. In the case of the leaning top, a torque is continually adding an angular velocity at right angles to the angular velocity of the top, with the result that the axis of the top (the direction of the angular velocity) is continually changing in direction.

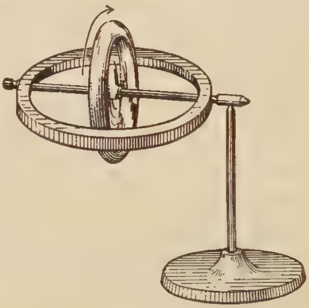


FIG. 100

The common toy gyroscope shows an interesting case of precession. Fig. 100 shows one form of the gyroscope. A gyroscope usually has a disk which is loaded at the rim, so that considerable kinetic energy can be given it. The disk is set in rapid rotation, and the gyroscope is then placed on the pivot, as shown in the figure. Instead of falling, it rotates in a horizontal circle. In Fig. 101 are shown the vectors which represent the solution of the problem. This figure is one in three dimensions; we must imagine that the lines are not all in the same plane. The vector OA, which represents the angular velocity of the disk, is drawn parallel to the axis of rotation of the disk. OZ, which represents the rotation added in a small time by the

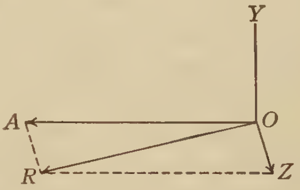


FIG. 101

torque due to the weight of the gyroscope, lies in a horizontal plane at right angles to *OA* and to the vertical line *OY*. The resultant of these two is the vector *OR*. Since both *OA* and *OZ* are in the horizontal plane, the resultant *OR* will also be in the same horizontal plane. The vector *OR* gives the direction of the axis of the resultant rotation; hence the line *OA* will have to move so that it will point in this direction. The action is really continuous; for as the axis moves toward the position *OR*, the vector *OZ*, which is always perpendicular to *OA*, will also move. The result is a precession: the gyroscope moves in a horizontal circle, in a counterclockwise direction as seen from above.

The gyroscope has many practical uses. It is used to steer torpedoes, to hold them to a direct course. It is used to stabilize ships; that is, to prevent them from rolling. One form of it has been developed to serve as a nonmagnetic compass, an apparatus which when set in rotation will always keep its axis of rotation parallel to the axis of the earth, even when its support is moved around.

The action which has been described above (that is, the tendency of the axis to precess) is shown in a number of cases. A side-wheel steamer when turning in a sharp curve has a tendency to tip sidewise, owing to the precession of the paddle wheels. The flywheel of the engine of an automobile tends to precess when the car is run around a curve. In the automobile the axis of the flywheel is parallel to the length of the car. In going around a curve the precession exerts a tendency to lift either the rear or front wheels, depending on which way the car is being turned. The tendency of rotating machinery, such as dynamos and motors, to precess on board a ship must often be taken into account in the design of the bearings and mountings of the machines.

Another interesting case of precession is that of the earth. If there were no torque changing the angular velocity of the earth, its axis would always point in the same direction. Owing to the fact that it is not a perfect sphere, the attraction of the moon tends to rotate the earth about an axis perpendicular to its axis of rotation. This causes the axis of the earth to precess slowly, the pole of the earth describing a circle which takes about twenty-five thousand years to complete. The result of this precession is a very slow shift of the seasons.

130. Resolution of angular velocities; components of rotation of the earth. Vectors which represent angular velocities may be resolved into components in the same way as other vectors. In Fig. 102, OA represents a rotation about the axis OX. The component of this rotation about the line OX' is OB. As an example take the case of the rotation of the earth. Let O be taken at the center, OX be the direction of the earth's axis, and OX' a line passing through the center of the earth and an observer. If the observer is at any point on the earth's surface except on the equator, the earth has a component of rotation about the line OX'. It is literally true that the earth is rotating about each one of us. The angular speed at which it is rotating can be found by measuring the magnitude of the component in a figure which is similar to Fig. 102, and which is drawn to scale.

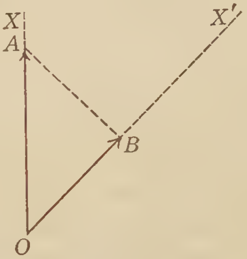


FIG. 102

One can see that the earth is really rotating about him if he thinks of it in the following way: If he faces south, he knows that points in front of him have a greater velocity* than he has, for they are nearer the equator, where the speed is greatest. They are then, relative to him, moving to his left, since the earth rotates eastward. If he faces north, he knows that he has a greater velocity eastward than points lying to the north; hence points to the north are moving relatively to his left. When he faces east, and thinks of the motion of the points to the east, where they will be in a few hours, he will see that these points are changing their direction of motion and are moving off to his left. For in about twelve hours from now the points in front of the observer will be on the other side of the earth's axis, but obliquely down to his left. In the same way, when he faces west the points lying westward have a component of motion to his left. This will be better understood if the student will try to point to places in Europe and Asia which are in the same latitude as he. The use of a globe or a tennis ball will help.

---

* This applies to an observer in the Northern Hemisphere.

Foucault devised an interesting experiment to show that the earth has a component of rotation around the observer. If a pendulum could be swung at the north pole and started vibrating in a plane, it would continue to vibrate in the plane. But the earth is rotating around the pole, and would rotate under the pendulum. To an observer the earth would appear to be at rest, and the plane of vibration of the pendulum would appear to rotate, making one complete rotation in twenty-four hours. At a point on the earth away from the pole the component of rotation is less, as seen in Fig. 102. Foucault swung a long pendulum, and found that with respect to the earth the plane of vibration rotated as expected. This experiment attracted great attention. Through the favor of Napoleon III the experiment was publicly repeated in the Pantheon, with a ball of 28 kilos suspended by a wire 67 meters long. The experiment was again repeated at the Universal Exposition, in 1855. It has since been performed in many laboratories.

131. Effect of the rotation of the earth on the motion of projectiles and winds. The rotation of the earth has an interesting and important effect on the motion of projectiles and winds. In the last section it was shown that the earth is actually rotating around an observer situated at any point on the surface of the earth except on the equator. This rotation is such that if the observer faces in any direction, all points in front of him are moving to his left. Hence, when he fires a gun at a distant target, he is really shooting at a moving object. In the absence of other deflecting forces the shell passes to the right of the point at which it is aimed. Hence, on account of the rotation of the earth, a projectile apparently tends to drift to the right.

This tendency of objects moving on the earth's surface to drift to the right is very marked in the case of the winds. For example, the trade winds—those surface winds which tend to move toward the heated regions of the equator—are deflected to the right, so that in the Northern Hemisphere they blow toward the southwest instead of due south. This tendency for a wind to be deflected toward the right is evident in other cases. This point will be discussed more fully in Chapter XXII.

132. Kinetic energy of rotating bodies; moment of inertia. The computation of linear kinetic energy is very simple, for kinetic energy is given by formulas (44) and (45):

$$\text{Kinetic energy} = \tfrac{1}{2}mv^2 \text{ (C.G.S. units)}$$

or

$$\frac{1}{2}\frac{w}{32}v^2 \text{ (B.E. units).}$$

But in a rotating body the speed v is not the same for all points of the body. It is far more convenient to use the angular speed ω, which is the same for all parts of the body. In that case the kinetic energy of rotation is found to be equal to $\frac{1}{2}\,I\omega^2$, where I is the constant called the *moment of inertia*. This constant depends on the mass of the body and on the distribution of the mass with respect to the axis. The expression for the kinetic energy in terms of angular speed can be derived from equation (44) as follows:

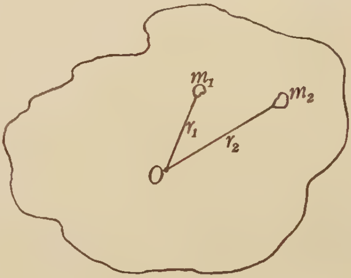


FIG. 103

Let the body which is rotating about O (Fig. 103) be the one for which the kinetic energy is to be determined. Let us in imagination divide this body into a large number of small parts. The total kinetic energy of the body can then be expressed by

$$\text{K.E.} = \tfrac{1}{2}m_1v_1^2 + \tfrac{1}{2}m_2v_2^2 + \tfrac{1}{2}m_3v_3^2 + \tfrac{1}{2}m_4v_4^2 + \cdots, \qquad (70)$$

where m_1 and v_1 are the mass and linear velocity, respectively, of the first part, m_2 and v_2 those of the second part, m_3 and v_3 those of the third part, and so on. From equation (69) we have

$$v_1 = \omega r_1,\ v_2 = \omega r_2,\ v_3 = \omega r_3,\ \cdots,$$

where ω is the angular velocity of the body in radians per second, and where r_1, r_2, and r_3 are the distances of m_1, m_2, m_3 from the axis of rotation. Substituting these values in equation (70), we have

$$\text{K.E.} = \tfrac{1}{2} m_1 \omega^2 r_1^2 + \tfrac{1}{2} m_2 \omega^2 r_2^2 + \tfrac{1}{2} m_3 \omega^2 r_3{}^2 + \cdots$$
$$= \tfrac{1}{2}\omega^2 (m_1 r_1^2 + m_2 r_2^2 + m_3 r_3^2 + \cdots).$$

The quantity in the parentheses is a *constant* of the body, being independent of the speed, depending only on the mass of the body and on the distribution of this mass with respect to the axis. This constant of the body is called the *moment of inertia*, and is usually represented by the letter I. Hence

$$I = m_1 r_1^2 + m_2 r_2^2 + m_3 r_3^2 + \cdots. \tag{71}$$

Using this in the preceding equation the kinetic energy becomes

$$\text{K.E.} = \tfrac{1}{2} I \omega^2. \tag{72}$$

This equation is true for both B.E. and C.G.S. units. Whenever the moment of inertia of a body and the angular speed ω are known, the kinetic energy of rotation is easily computed.

Instead of stating the angular speed in radians per second, it can be expressed in revolutions per second.

Since $$\omega = 2\pi N,$$

where N is the number of revolutions per second,

$$\text{K.E.} = \tfrac{1}{2} I (2\pi N)^2,$$

or $$= 2\pi^2 N^2 I. \tag{73}$$

The moment of inertia of many regularly shaped bodies can be computed mathematically, but for irregular bodies it must be determined by experiment. The following are some of the formulas for the moments of inertia of regularly shaped bodies; in each case m is the total mass of the body.

Circular disk; axis a diameter of the disk, radius $= r$:

$$I = \tfrac{1}{4} m r^2$$

Solid cylinder; axis through the center perpendicular to the axis of cylinder, length $= l$, radius $= r$.

$$I = m\left(\frac{l^2}{12} + \frac{r^2}{4}\right)$$

Solid cylinder or disk; axis coincides with the axis of the figure, radius $= r$:

$$I = \tfrac{1}{2} mr^2$$

Sphere; radius $= r$, axis any diameter:

$$I = \tfrac{2}{5} mr^2$$

When B.E. units are used, weight in pounds divided by 32 is substituted for the mass m.

133. Radius of gyration. If R^2 is a mean value of r^2 such that

$$mR^2 = m_1 r_1^2 + m_2 r_2^2 + m_3 r_3^2 + \cdots,$$

then

$$I = mR^2,$$

where m is the total mass (see equation (71)). R in this expression is called the *radius of gyration.*

134. Formulas analogous to those of linear motion. In *linear* motion, by (21),

$$F = ma,$$

In *rotational* motion

$$L = I\alpha, \tag{74}$$

where L is the moment of force, or torque, and α the angular acceleration. In *linear* motion we have, by (26), as a definition of momentum,

$$\text{Momentum} = mv;$$

while, in *rotational* motion,

$$\text{Angular momentum} = I\omega. \tag{75}$$

The kinetic energy for *linear* motion is, by (44),

$$\text{K.E.} = \tfrac{1}{2} mv^2,$$

and for *rotational* motion, by (72),

$$\text{K.E.} = \tfrac{1}{2} I\omega^2.$$

In equation (75) angular momentum is given as the product of the moment of inertia times the angular velocity. In the case of a mass such as a small stone whirling on the end of a string, where all points of the mass have practically the same distance from the center, a simpler definition can be given. For this simple case the product of the *linear momentum* times the *radius* of the rotational motion is equal to angular momentum, or moment of momentum, as it is sometimes called.

135. Conservation of angular momentum. Newton's first law of motion may be stated thus: *A body maintains a constant linear momentum unless acted upon by some external force.* Similarly, it is true that a rotating body tends to maintain a *constant angular momentum* unless acted upon by some external torque.

Since angular momentum is a vector quantity, to say that the angular momentum tends to remain constant means not only constant in magnitude but also constant in direction. For example, in revolving around the sun the earth tends to keep the angular momentum about its own axis constant. Hence the earth's axis always points in the same direction,* approximately toward the North Star. A man throwing a quoit or a horseshoe, and wishing to prevent its turning over while in flight, gives it a spin. It will then travel with the axis of rotation always pointing in the same direction. In the same way, a coin when given a spin about an axis perpendicular to its face will travel without turning over. The barrel of a rifle or a cannon is "rifled" to give the projectile a rotation. A rotating projectile will travel without "tumbling" and will strike an object with its nose. The same thing is beautifully shown by a kicked or thrown spinning football, often called a *spiral.*

There are a large number of simple cases which are readily explained by this principle, which states that angular momentum tends to remain constant. Take the case of a stone on the end of a string, the other end of the string being held by the right hand. Let the string and the stone be given a considerable angular velocity. If the string is drawn with the left hand through the right hand, so that the part rotating with the stone is shortened, the angular velocity of the stone and string will *increase.* This increase is due to the fact that the angular momentum tends to remain constant. Shortening the string decreases the *moment of inertia* of the stone. Since the angular momentum is the product of the moment of inertia and the angular velocity, the angular momentum can remain constant only in case the angular velocity increases at the same rate that the moment of inertia decreases.

---

* This is only approximately true, for there are small external moments of force acting on the earth. These produce a slow precession of the axis (sect. 129).

Nearly every child learns to "work up" in a swing, but few people know the principle involved. The child usually makes many unnecessary motions and usually does not know which are the essential ones. The necessary motions are (1) to raise the center of gravity of the body as the swing moves through the middle position, where it has the greatest velocity; (2) to lower the center of gravity of the body with respect to the swing at the extreme end of the motion. The action is explained as follows: When one raises his center of gravity at the middle part of the motion, where the speed is the greatest, he shortens the *radius* of the angular motion and hence decreases his moment of inertia. But in this part of the motion the resultant of the external forces is zero; hence the angular momentum is constant. If the angular momentum is not to change as the swing moves through the central part, the angular velocity must increase when the moment of inertia is decreased. At the end of the swing, where the velocity is zero, the angular momentum is also zero, and the lengthening of the radius by lowering the center of gravity does not affect the action.* (If one lowers his center of gravity at the middle point and raises it at the end of the swing, the motion is quickly arrested. Why?)

When water which is escaping from the central outlet in the bottom of a washbowl is given a slight rotational motion, the angular velocity of the water as it nears the center increases. This is due to the fact that the angular momentum of the water tends to remain constant, and this can be done for shortening radii only by an increasing angular velocity. In this case the action of the centrifugal force of the water is also clearly shown. It has the effect of an outward pressure banking the water up, so that often there is a central cavity down through the moving mass of water.

The motion of tornadoes is somewhat similar to that of the water in the washbowl. In the center of the tornado there is a strong uprush of air, and the whirling mass of air is moving inward, just like the water in the bowl. The tendency for the angular momentum of the air to remain constant explains the enormous velocities which are frequently developed.

---

* When one raises himself as the swing moves with its maximum speed through the middle part, he does work against centrifugal force. This is the work that supplies the energy.

Tumbling in gymnasium work furnishes many fine examples of this tendency of angular momentum to remain constant. For example, when a man jumps off a springboard, he can give himself a certain amount of angular momentum. If he now doubles up, bringing his head, arms, knees, and feet near his center of gravity, he will decrease his moment of inertia and thus increase his angular velocity. He may increase his angular velocity enough so that he can turn more than one somersault while in the air. Just before he reaches the floor he straightens out, increasing his moment of inertia and thus decreasing his angular velocity.

136. Total kinetic energy. A moving body may have kinetic energy both on account of linear translational motion and on account of rotational motion. If a sphere rolls down an incline, the total kinetic energy developed is given by the formula

$$\text{K.E.} = \tfrac{1}{2}\,mv^2 + \tfrac{1}{2}\,I\omega^2,$$

where v is the linear velocity of the center of the sphere, ω the angular velocity about a line through the center, m the total mass, and I the moment of inertia about a line through the center.

A sphere and a cylinder of equal masses do not have the same moment of inertia. They will therefore roll down an incline with different linear speeds. (Why?) Not all the energy given a rifle or a cannon ball goes into linear kinetic energy; some of it goes into rotational energy. The same is usually true of a thrown or batted baseball.

PROBLEMS

1. A circular disk (or cylinder) has a radius of 6 in. and weighs 10 lb. Compute its moment of inertia about an axis perpendicular to the disk.

2. How many foot-pounds of energy will the disk of Problem 1 have when making 4 revolutions per second?

3. How much work would be required to give a spherical ball, mass 4 kg. and radius 5 cm., an angular speed of 4 revolutions per second?

4. A flywheel which is a uniform circular disk of 12 in. radius weighs 320 lb. How much energy is stored in it when it makes 600 revolutions per minute?

5. A rod, the mass of which may be neglected, has a mass which weighs 4 lb. fastened on one end, and one which weighs 5 lb. on the other. The distance between the masses is 4 ft. Compute the moment of inertia of rod and weights when rotated about a line perpendicular to the middle point.

6. A light rod 30 cm. long has a mass of 1 kg. on each end. How much torque would be required to give this rod in 20 sec. an angular speed of 2 revolutions per second about an axis perpendicular to the middle point of the rod?

7. Two iron balls, weighing 25 lb. each, are connected by a very light rod so that their centers are 3 ft. apart. (*a*) Compute the moment of inertia about an axis perpendicular to the rod at its center. (*b*) How much torque would be required to give this an angular speed of 360 revolutions per minute in 20 sec.?

8. The entire weight, 200 lb., of a wheel 3 ft. in diameter is practically concentrated at the rim. (*a*) Compute its moment of inertia. (*b*) What torque will be necessary to give it a speed of 1 revolution per second in 5 sec.?

CHAPTER XII

GRAVITATION

Gravitation, 137. Gravitational attraction is universal, 138. Newton's law of gravitation, 139. Mass of the earth, 140. Value of the gravitational constant, 141. Variations in the value of *g*, 142. Tides, 143. Cause of tides, 144.

137. Gravitation. Nearly everyone knows that the reason a stone falls to the ground is that the earth attracts the stone. There are, however, things about this which not all know. Is this attractive power possessed only by the earth? Do all masses attract each other? Will a plumb line hung near a mountain be pulled out of its vertical position by the attraction of the mountain? Is Newton's third law of motion true for this attraction? That is, if the earth attracts a stone, will the stone attract the earth? Everyone at some time or other wonders about the mysterious forces which hold the moon and the planets in the sky. Can these be explained? Largely through the aid of the imagination and the keen mind of Sir Isaac Newton these questions have been answered.

138. Gravitational attraction is universal. Newton was the first to see that the attraction of the earth for bodies on its surface was only a special case of a general law. According to his idea this attraction was not something possessed by the earth alone but was common to all bodies; every body attracted every other body, it made no difference how large nor how small.

Bouguer, in Peru, about 1740, was the first to show the attractions of masses smaller than the earth. He found that a plumb line hung near a mountain was not parallel to another hung at some distance from the mountain. This deviation was due to the attraction between the plumb bob and the mountain. A later modification was to use two plumb lines hung on opposite sides of the mountain. In all cases the deflection of the plumb lines was very small and required very careful measurement to detect.

Several different types of experiments with sensitive balances have been devised. In a method due to Poynting two equal lead balls, of about 50 pounds each, were suspended from the two ends of the beam of a balance. A third mass of lead, of about 350 pounds, was arranged so that it could be moved from directly under one of the 50-pound balls to a point directly under the other ball. The attraction of the large mass produced a change in the apparent weight of the suspended balls. The observed change was equal to that obtained when a mass of four tenths of a milligram was added to one side of the balance. The effect was small but measurable. The attraction between smaller masses than these was shown by an extremely delicate experiment designed by Professor Boys. He was able not only to show the existence of attractive forces between small metallic spheres but to measure the forces accurately. Newton's supposition that the attraction the earth exerted is only a special case has been fully verified.

139. Newton's law of gravitation. *Every particle in the universe attracts every other particle with a force directly proportional to the product of the two masses and inversely as the square of the distance between them.*[*]

Stated algebraically,
$$F = G\frac{m_1 m_2}{d^2}, \qquad (76)$$

where G is a proportionality factor the numerical value of which depends on the units used in measuring the other quantities. It is usually called the *constant of gravitation.*

Newton did not deduce this law by a rigorous process. He simply made an intelligent guess and then proceeded to verify it by applying it to the motion of the moon around the earth. This method of procedure is the common one in experimental scientific work. From known facts one guesses at the existence of others, or one supposes that there must be a general law of which an observed fact is a special case. The guess, or hypothesis, must be put to an experimental test. Only after being verified in a large number of cases is it accepted.

---

[*] According to Einstein's *theory of relativity* Newton's law of gravitation is not exact but an extremely close approximation.

Newton's test of this law can be explained in the following way: Applying the law to the moon, the force acting on the moon is $G\frac{Mm}{r^2}$, where M is the mass of the earth, m the mass of the moon, and r the distance between their centers. The acceleration given to the moon will be the force divided by the mass of the moon (see equation (21)), or $G\frac{M}{r^2}$. The moon must be falling toward the earth with this acceleration. However, we have seen in section 77 that when a body is moving in a curved path there will always be an acceleration toward the center of the curve. Furthermore, we found that this acceleration must always be equal to v^2/r, where v is the linear speed and r the radius of the curved path. It is known that the moon revolves in an approximately circular orbit around the earth. Hence it must have an acceleration equal to v^2/r. In the case of the moon v and r are known; hence the acceleration can be computed. The next step is to see if this computed acceleration agrees numerically with that computed from the law of gravitation; that is, if it is equal to $G\frac{M}{r^2}$. The value of the product GM can easily be determined. A mass m' on the surface of the earth is attracted by the force $G\frac{Mm'}{R^2}$, where R is the radius of the earth. But this force is the weight of the mass m', or $m'g$ (see equation (22)).

Hence
$$G\frac{Mm'}{R^2}=m'g,$$
or
$$GM=gR^2.$$

Since both the quantities on the right are known, the product GM is readily obtained, and hence the acceleration GM/r^2 can be computed. This is found to agree closely with the observed acceleration v^2/r.

Further confirmation of this law is obtained by applying it to planets. Not only is the general motion of a planet predicted, but perturbations, or disturbances of its orbit due to the approach of other planets, are accurately forecast. The accuracy with which an eclipse can be predicted is an example of an experimental test of the law of gravitation.

140. Mass of the earth. If the gravitational constant G is known, it is a simple matter to determine the total mass of the earth. One gram on the surface of the earth is, by the law of gravitation, attracted by a force

$$F = G\frac{M}{R^2}.$$

This equation is obtained from equation (76) by setting one of the masses equal to unity and the other to M, the mass of the earth. R is the radius of the earth. But this force is the weight of 1 gram, or, approximately, 980 dynes.

Hence
$$980 = G\frac{M}{R^2}.$$

G has been found to be equal, approximately, to 6.6×10^{-8} (see section 141), and R equal to 6×10^8 centimeters. Computation shows that M is equal to 5.3×10^{27} grams, approximately. This, when worked out with greater numerical precision, is the best method for determining the mass of the earth.

Since the volume of the earth is known, the mean density, or the mass divided by the volume, is easily found. The mean density as found in this way is 5.53. As the average density of soils, rocks, etc. near the earth's surface is less than half of this, the density of the earth must be much higher in the interior than near the surface.

We are now able to answer another common question. The motions of celestial bodies, and the observed attraction between masses in laboratory experiments, show that gravitational force obeys the ordinary laws of mechanics. For example, when the earth pulls on a stone, the stone pulls on the earth; action is equal to reaction. If this is true, why is not the earth accelerated toward a falling stone at the same time that the stone is accelerated toward the earth? We must answer that it probably is. But the effect on the earth is immeasurably small. We can compute the acceleration that the earth will have. Let the mass of the falling body be 1 kilogram. The force acting on the earth and that acting on the kilogram are equal. In section 63 we saw that when the same force acts on different masses,

$$\frac{a_1}{a_2} = \frac{m_2}{m_1}.$$

In this case m_1 is the mass of the earth, m_2 the mass of the kilogram, a_1 the acceleration of the earth, and a_2 the acceleration of the kilogram mass, or 980 centimeters per second per second. From the above

$$a_1 = 980 \frac{1000}{5.3 \times 10^{27}} = \frac{1.85}{10^{22}} \text{ cm./sec.}^2$$

This acceleration is so small that if it lasted for a hundred years the velocity acquired by the earth would be too small to measure!

141. Value of the gravitational constant. The gravitational constant G is found by carefully measuring the very minute force of attraction between two known masses at a known distance apart. G is then computed by the aid of equation (76). The value found by Professor Boys is, when all quantities are measured in C.G.S. units,

$$G = 6.6576 \times 10^{-8}.$$

142. Variations in the value of *g*. In Chapter X it was pointed out that the value of g could be found with the aid of a pendulum. It is found that the value of g is not the same at different places on the earth.

ACCELERATION OF GRAVITY*

PLACE	ELEVATION IN METERS	g CM./SEC.^{2}
Ithaca, New York	247	980.300
Washington, District of Columbia	103	980.095
Beaufort, North Carolina	1	979.729
Cleveland, Ohio	210	980.241
Madison, Wisconsin	270	980.365
St. Louis	154	980.001
New Orleans	2	979.324
Galveston, Texas	3	979.272
Denver	1638	979.609
Pikes Peak	4293	978.954
Seattle	58	980.733
San Francisco	114	979.965

*From the Smithsonian Physical Tables.

Several things tend to make g different at different places: (1) At higher elevations the gravitational force is less, and hence the acceleration g must be less. (2) The observed weight of any body is its true weight less the centripetal force required to keep it rotating with the earth. Since the weight is less, g must also be less. The centripetal force is greatest at the equator and least at the polar regions. It is greater at high elevations than at sea level. (3) Apparently the earth is not of uniform density. Hence at those places where denser materials lie near the surface the gravitational attraction is greater than normal.

143. Tides. To anyone who has spent much time at the seashore the tides are a source of unfailing interest. The clocklike regularity and the enormous masses of water moved create strong impressions. The water rises to high tide, then falls to low tide, then returns to high tide again, and so on with a ceaseless ebb and flow. The interval between two high tides is about twelve hours and twenty-five minutes. The height of the tide in mid-ocean is from 2 to 5 feet, but along the coasts it rises much higher. On the Atlantic coast of the United States the tide usually rises from 6 to 12 feet, varying from place to place. Farther north on the Atlantic coast the rise is very high; for example, in the Bay of Fundy it sometimes rises 70 feet. On the English and French coasts it occasionally rises as high as 40 or 50 feet.

When one knows that the time interval between low and high tide is only a little over six hours, and that in that time the level of the water in our harbors and bays must change a considerable amount, then he can understand why the rise and fall of the tide must be accompanied by such strong currents. In the entrances of many harbors and bays the current is often so strong that only high-speed boats can make headway against it.

144. Cause of tides. It is not true that the moon revolves around a stationary earth. The force of attraction which the earth exerts on the moon causes the moon to move in a curved path. But what does the force which the moon exerts on the earth do to the earth? Certainly the earth cannot remain stationary; it must be accelerated by this force. The answer is that both the earth and the moon revolve around the axis XO of Fig. 104 with the same angular speed. The period of one complete rotation is about twenty-seven days. The axis runs through a point called the center of mass of the earth and moon and lies just inside the earth's surface. By referring to this figure, and keeping in mind that the earth and the moon are whirling around this axis, one can understand why there is a tendency of the ocean water to move toward A. It is what we have called the centrifugal action.

At *B* the centrifugal force is less than at *A*, but there the direct attraction of the moon tends to lift the water. The result is that there is a reduction of the weight of water at both *A* and *B*, and hence a tendency to produce tides at those places.*

In addition to the motion already described, the earth has a rotation around the axis *MN* with a period of twenty-four hours. This is the motion which gives us day and night. The earth spinning on the axis *MN* carries through the regions marked *A* and *B* all parts of the surface of the earth which are not near the poles. The surface of the earth is thus forced through the tides; or, as they appear to those who are moving with the earth, the tides travel around the earth as extremely long waves. The continents and other masses of land interfere with the free motion of these waves, so that the actual direction and magnitude of the wave-motion become very complicated.

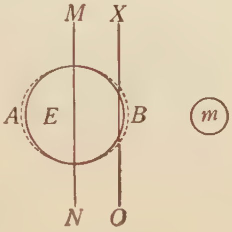


FIG. 104

The sun produces an effect, but a much smaller one than that produced by the moon; in fact, an effect so small that it is usually masked by the tide due to the moon. At both new and full moon the tide-producing forces of the moon and the sun act together, producing larger tides, which are called spring tides. This condition occurs twice a month.

Since the cause of the tidal flow is the rotation of the earth about the axis *MN*, the source of energy in the tides is the kinetic energy of rotation of the earth. Friction in the ebb and flow of the tides should tend to slow up the earth's rotation.

PROBLEMS

1. What would a body which weighed 100 lb. on the earth's surface weigh when placed at a distance equal to the earth's radius above the surface?

2. Compute the force of attraction in dynes between two spheres of 5 kg. each which are placed with their centers 20 cm. apart.

3. What is the ratio of the weight of a kilogram on the surface of the earth to what it would be if raised 10 km. from the surface? (Assume the radius of the earth to be 6.3×10^8 cm.)

---

* The attraction the moon exerts on the water at *A* decreases the effect of the rotation about the axis *OX*, but the necessary centripetal force is much greater than the attraction of the moon; hence water has less weight at *A*. Computations show that the losses of weight of unit mass of water at *A* and *B* are approximately equal.

4. Find the ratio of the weight of a body in Madison, Wisconsin, to what it would be in New Orleans (see section 142).

5. What would be the acceleration of a freely falling body if the body is placed 1000 mi. above the earth's surface? (Assume the diameter of the earth to be 7920 mi. and g, at the surface, to be 32.2 ft./sec.^{2})

6. How many pounds will a man who weighs 150 lb. on the earth weigh on the surface of the moon? (The mass of the moon is 0.0128 of that of the earth, and the radius of the moon is 0.273 of the earth's radius.)

7. A lead ball, the mass of which is equal to 10 kg., is suspended from the arm of a balance. (*a*) What will be the attractive force on it when a 200-kilogram ball is placed directly under it, the distance between centers being 25 cm.? (*b*) What will be the apparent increase in the mass of the suspended ball?

CHAPTER XIII

FLUIDS IN MOTION

Loss of pressure due to fluid friction, 145. Change in pressure produced when speed of liquid is changed, 146. Explanation of decrease in pressure with increased speeds, 147. Illustrations of decrease of pressure with increase in speed, 148. Bernoulli's theorem, 149. Proof of Bernoulli's theorem, 150. Summary, 151. Energy of a liquid, 152. Torricelli's theorem, 153. Transference of hydraulic energy of water to machines, 154. Water wheels and turbines, 155. Hydraulic ram, 156. Pitot tube, 157.

Any understanding of the simple properties of fluids that does not include facts regarding fluids in motion is incomplete. Some of the facts which are true for liquids at rest are not always true for moving liquids. An example in point is the theorem which states that at all points on the same level in connected vessels the pressure is the same (sect. 23).

The use of water power has become a matter of great economic importance; hence it is of considerable interest to know something about the principles of moving liquids.

145. Loss of pressure due to fluid friction. The most noticeable effect of fluid friction is in the loss of pressure in pipe lines. Force must be used to overcome the retarding forces of friction. Hence in any pipe line where water or any other liquid is flowing, there is a continuous decrease in pressure as the distance from the source of supply increases.

In pipes of the same size the loss of pressure on account of friction is proportional to the length of pipe and approximately proportional to the velocity of the fluid. The loss of pressure in large pipes is less than in smaller ones, provided the velocity of flow is the same in the two.

146. Change in pressure produced when speed of liquid is changed. Fig. 105 shows diagrammatically an experiment in which the water is flowing through a constriction in a pipe. The liquid has a greater speed at B than at either A or C. The height of the water in the vertical tubes shows that the pressure in the pipe is greater at A and C than at B. The pressure at C will be somewhat less than that at A on account of the fluid friction. But the large decrease in pressure from A to B and the rise of pressure from B to C cannot be due to any form of fluid friction.

This experiment shows what is found to be a general fact: in a fluid which is in motion there is *a decrease in pressure wherever the speed is increased, and an increase in pressure wherever the speed is decreased.*

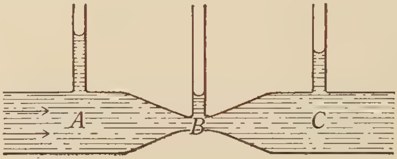


FIG. 105

147. Explanation of decrease in pressure with increased speeds. The pressure of the liquid at B of Fig. 105 is transmitted through the liquid at A from the pump or source of pressure. But the liquid between A and B is being *accelerated.* No body which is being accelerated transmits all the force or pressure which is applied to it. Take the simple case of a book lying in one's hand. The earth exerts a downward force on the book. This is transmitted to the hand, producing an equal downward force on it. But if the hand and the book are accelerated downward, not all the force due to the weight of the book is transmitted to the hand: the pressure on the hand is decreased because part of the force is used in accelerating the book. So it is that the pressure transmitted to B is less than at A, because part of it is used in accelerating the water between A and B. In the space between B and C there is a negative acceleration, and the pressure is greater at C than at B.

148. Illustrations of decrease of pressure with increase in speed. Let A be a point in the water of a reservoir, and B a point in the outlet pipe (Fig. 106). Although the points A and B lie on the same level, the pressures will not be the same at these points, for at B the speed of the liquid is much greater than at A. As explained in the last section, this is due to the fact that the water between A and B is being accelerated. The observed decrease in pressure in the water pipes in a house when water is flowing from one of the faucets is due partly to this effect.

Two light balls (for example, hollow celluloid balls) are suspended by strings so that they hang two or three inches apart. If one blows in between the balls, they will move together at right angles to the direction of the blast. In Fig. 107 the direction of the motion of the blast of air is shown by the lines *AA′*. Between the balls the air has considerable speed, while beyond, at *A′*, where the blast spreads out, the speed is much less. It was shown in connection with Fig. 105 that the pressure at *B* must be less than at *C*. For a similar reason the pressure between the balls must be less than at *A′*. At *A′*, or a little beyond, where the air comes to rest, the pressure must be the normal pressure of the air. Hence the pressure between the balls is less than atmospheric pressure, and they are pushed together by the air pressure on the outside.

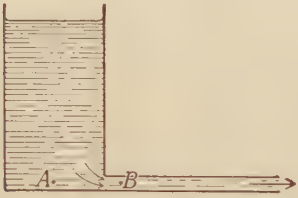


FIG. 106

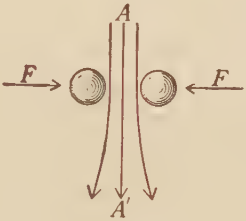


FIG. 107

A similar thing happens when two boats are lying side by side in a current of water. It is a well-established fact that observed phenomena depend only on *relative* motion and not on *absolute* motion. Since there are forces pushing the boats together when they are in moving water, it follows that if they are moving side by side in stationary water there will also be forces pushing them together. For the relative motion of the boats and the water is the same in the two cases. Serious collisions have been caused when one boat in trying to pass another has come close enough to be affected by the apparent attraction. When airplanes began flying in squadrons, several wrecks occurred which were due probably to this cause.

A light ball, such as one made from cork, can be balanced in an air jet, or a much heavier ball on a water jet. If the ball should move off to the side (Fig. 108), the air on one side will have a greater speed than on the other. There will then be a lower pressure on one side than on the other. Hence the ball will be forced back into the jet. The ball may oscillate from one side of the jet to the other, but under proper conditions it will rarely leave the jet.

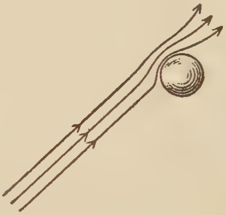


FIG. 108

A thin cork with a central hole is mounted tightly on the end of a glass tube (an ordinary spool will do), and a card is held loosely in place by a pin, as shown in Fig. 109. If one blows through the tube, the card is not blown off but is drawn up tightly to the cork. The air spreading out into the narrow space between the card and the cork has a greater speed than the air outside, with the result that the pressure of the air between the card and the cork is below atmospheric pressure. "Ball" nozzles work on the same principle.

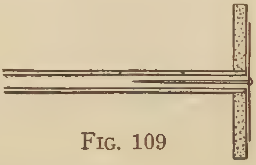


FIG. 109

Suppose that a baseball (Fig. 110) is moving toward the right, and at the same time rotating as indicated. A moving ball piles up in front of it a cushion of air. But if the ball is rotating, air is swept around by the motion of the ball. At the side *B* the air is moving into the cushion, where the pressure is greater; hence it does not acquire a high speed. But at *A* the air is being carried away from the front, where the pressure is high, to the rear, where the pressure is lower. Thus at *A* the air acquires a high speed, with a lowering of the pressure. Since the pressure is greater at *B* than at *A*, the ball is given an acceleration to one side (up, in the figure). In addition to this effect, the rotation prevents a symmetrical piling up of the air in front of the ball. Air is being swept away on one side and brought forward on the other. There will thus be more of a cushion on that part of the

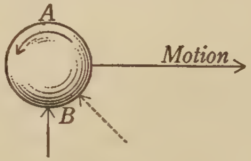


FIG. 110

front toward *B* than toward *A*. This cushion does not, then, push straight back, but in the direction shown in the figure by the broken arrow. This force has a component tending to push the ball out of a straight path (upward, in the figure). These are the forces that produce the curve of a thrown baseball or a driven tennis or golf ball. These forces will change in direction when the direction of the rotation of the ball is changed. Hence a pitcher can control the motion of a baseball by the rotation he imparts to it. He also knows that to obtain a wide curve a very high rotational speed must be given the ball.

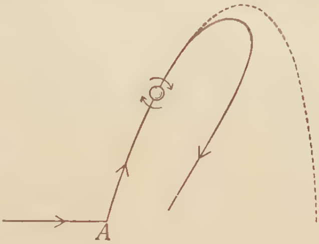


Fig. 111

There is a case of the curving of a baseball which can be easily seen by spectators. It is a high foul going nearly straight up. The foul is produced when the bat strikes the lower side of the ball. In addition to changing the direction of the motion of the ball, the bat imparts a rapid rotational motion to the ball. Instead of following the path of the broken line in Fig. 111, the ball curves as shown by the full line. One often sees a catcher run toward the stand for one of these fouls and then be compelled to run back rapidly toward the plate to catch it. If one flips up a light celluloid ball, giving it a rapid rotation with his fingers as it leaves his hand, it will describe a similar path.

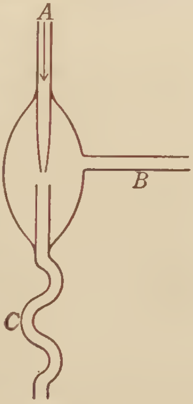


Fig. 112

Fig. 112 shows a simplified sectional view of an aspirator, or filter pump. The pressure of the water which enters at *A* is reduced below atmospheric pressure at the contraction in the jet, and air or water will flow in at *B*. Aspirators connected to city water mains will produce a partial vacuum of about one twentieth of an atmosphere. For this reason they are commonly used in physical and chemical laboratories as air pumps. They are also used to produce a partial vacuum in the automatic vacuum-brake systems and in the condensers of steam turbines. Sometimes they are used to pump water. The steam injector is a jet pump used to supply water to the boiler. In this case steam from the boilers is supplied at A, and water is drawn in at B. The common atomizer and some forms of spray pumps work on a similar principle.

149. Bernoulli's theorem. The *quantitative* law with regard to changes in pressure with changes in speed is due to Bernoulli. The theorem bearing his name states that at any two points in the same horizontal plane in a fluid,

$$\tfrac{1}{2}mv_1^2 + p_1 = \tfrac{1}{2}mv_2^2 + p_2, \tag{77}$$

where m is the mass of unit volume of the fluid, v_1 and p_1 the speed and pressure at one point, and v_2 and p_2 the speed and pressure at the other point.

An equivalent statement of this theorem is as follows: As the fluid flows from one place to another the quantities v and p change in such a manner that the expression

$$\tfrac{1}{2}mv^2 + p$$

remains constant. When the speed v increases, the pressure decreases. This is in agreement with the experiments described in sections 146 and 148. However, this gives more information, for it enables one to compute how much the change in pressure is when the speeds are known.

When there is no change in the speed of the liquid, the expressions containing v will cancel out of equation (77). It then reduces to a statement that

$$p_1 = p_2,$$

the theorem, given in section 23 for liquids at rest, that the pressure is the same at all points in a horizontal plane.

As will appear in the proof of this theorem, *no allowance is made for a loss of pressure due to friction*; hence the observed pressure changes are usually a little different from those given by Bernoulli's theorem.

If the flow is not in a horizontal direction, then the pressure will be changed because of changes in elevation. This change in pressure has been shown (see section 24) to be equal to hdg, in C.G.S. units, where g is the acceleration of a freely falling body, d the density of the liquid, and h the increase in height. For this case Bernoulli's theorem may be written thus:

$$\tfrac{1}{2}\, dv^2 + p + hdg = \text{a constant}, \tag{78}$$

for the mass per unit volume, m, is equal, when C.G.S. units are used, to the density d. For water, in C.G.S. units, the expression is

$$\tfrac{1}{2}\, v^2 + p + 980\, h = \text{a constant}. \tag{79}$$

In B.E. units the mass of unit volume of water is $\frac{62.4}{32}$ slugs (see section 66). Hence, for the case of water,

$$\frac{1}{2}\,\frac{62.4}{32}\,v^2 + p + 62.4\, h = \text{a constant},$$

or

$$0.97\, v^2 + p + 62.4\, h = \text{a constant}, \tag{80}$$

where p is measured in pounds per square *foot*, v in feet per second, and h in feet.

150. Proof of Bernoulli's theorem. Let p_1 be the pressure at A in Fig. 113, p_2 the pressure at B, and s the distance between A and B. The difference in the forces on the ends of a column of liquid 1 square centimeter in section and s centimeters long is $p_1 - p_2$. The average difference of the force on the opposite sides of 1 cubic centimeter of this column is

$$\frac{p_1 - p_2}{s}.$$

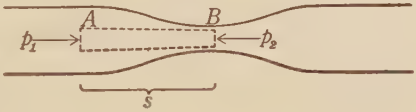


FIG. 113

This is the average force tending to accelerate 1 cubic centimeter of the column toward B. The work done by this force in moving the cubic centimeter from A to B is equal to the force times the distance, or

$$\frac{p_1 - p_2}{s}\, s = p_1 - p_2.$$

This work is equal to the *increase* in the kinetic energy of the cubic centimeter. Hence

$$p_1 - p_2 = \tfrac{1}{2}\, m v_2{}^2 - \tfrac{1}{2}\, m v_1{}^2,$$

where v_1 is the speed at A, v_2 the speed at B, and m the mass of 1 cubic centimeter. From this, by a simple transposition,

$$\tfrac{1}{2}\, m v_1{}^2 + p_1 = \tfrac{1}{2}\, m v_2{}^2 + p_2.$$

This is the algebraic statement of Bernoulli's theorem, given in equation (77).

Bernoulli's theorem has many practical applications. It is often used to compute the speed of flow in pipes or in orifices. Since friction has been neglected, the result will not be exactly correct; but the results obtained from this theorem are sufficiently accurate for many types of problems. For example, the method of measurement of the velocity of flow by a Venturi water meter is based on this theorem.

151. Summary. There are three circumstances under which the pressure in flowing liquids is changed.

1. There are changes in pressure on account of friction.
2. There are changes when the speed of the liquid changes.
3. There are changes when the liquid flows from one level to another.

In general, the actual changes observed are a combination of all three of these cases, but there are many instances in which the change in pressure is due almost wholly to one cause.

152. Energy of a liquid. It is usual in discussing the energy of a liquid to refer to the energy of a unit volume—a cubic foot, or a cubic centimeter, or some other convenient unit. The kinetic energy of 1 cubic foot of water is (see equation (45) of section 88)

$$\frac{1}{2}\frac{62.4}{32}v^2 = 0.97\ v^2 \text{ foot-pounds,}$$

where v is measured in feet per second and 62.4 is the weight in pounds of 1 cubic foot. Or if the cubic centimeter, which has a mass of 1 gram, is taken as the unit of volume, the energy is

$$\tfrac{1}{2}v^2 \text{ ergs,}$$

where v is measured in centimeters per second. In general, the kinetic energy of a liquid the specific gravity of which is d is

$$0.97\ dv^2 \text{ foot-pounds per cubic foot,}$$

or

$$\tfrac{1}{2}\ dv^2 \text{ ergs per cubic centimeter.}$$

The potential energy of a liquid can be measured by the work done in raising the liquid to its position. The work done in raising a cubic foot of water, or 62.4 pounds, to a height h is

$$62.4\ h \text{ foot-pounds,}$$

where h is measured in feet; or, if the liquid is of specific gravity d,

$$62.4\ dh \text{ foot-pounds.}$$

In C.G.S. units the work done in raising a mass m of water is, since its weight is mg,

$$mgh = 980\ mh \text{ ergs},$$

where h is measured in centimeters. The potential energy of 1 cubic centimeter of water is

$$980\ h \text{ ergs}.$$

In computing the total gravitational energy in a reservoir the height used must be the height of the center of gravity of the water.

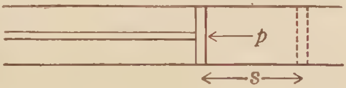


FIG. 114

The *work* done by a pump in forcing a *unit volume* of water into the city mains *is numerically equal to the pressure** in the mains if the pressure is kept constant. This may be proved as follows: If the piston of Fig. 114 is moved a distance s, a volume of water equal to As, where A is the area of the piston, will be forced into the mains.

The work done is

$$\text{Force} \times \text{distance} = pA \times s = p \times As,$$

where p is the pressure in the mains. Since the volume forced in is As, the work done per unit volume is equal to p. (If the cubic foot is used as a unit of volume, the pressure must be the force on a *square foot*.)

It follows that when water is drawn from a hydrant, the work done by the pump must be equal to the volume withdrawn times the constant pressure maintained at the pumping station. If the pressure is 40 pounds per square inch, or $40 \times 144 = 5760$ pounds per square foot, the work done for each cubic foot of water is 5760 foot-pounds. Since there are approximately 7.5 gallons in a cubic foot, the work done for each gallon at a pressure of 40 pounds per square inch is approximately 5760 divided by 7.5, or 770 foot-pounds. This neglects the work the pump must do in lifting water up to the level of the pump. The pumping of water requires more energy and is a more expensive process than is usually supposed.

---

*Some writers assume from this equivalence that a fluid has *pressure* energy. While in certain cases this is a convenient supposition, yet it is one which leads to many fallacies.

153. Torricelli's theorem. An old theorem due to Torricelli illustrates the convertibility of potential energy into kinetic energy. It may be stated as follows: *Water issuing from an orifice* (*B* in Fig. 115) *in a reservoir has, if all friction is neglected, sufficient speed to raise it to the level of the top of the water.* This theorem is easily seen to be true from the following consideration:

Suppose that water is allowed to flow out at *B* for only a short interval. What changes have taken place in the reservoir during that time? One notes that a layer of water is gone from the top and that the reservoir has lost energy by the amount of energy the missing layer had. The lost energy is equal to the amount of work necessary to raise an equal volume of water to the top. The water issuing from the jet must have the missing energy in the form of kinetic energy. It must therefore have enough to raise it to the level of the water in the reservoir.

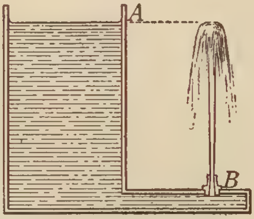


FIG. 115

The potential energy at the top is wh, where w is the weight of the water and h is the height above the orifice. The kinetic energy of the issuing jet is $\frac{1}{2}\frac{w}{g}v^2$. Since these must be equal,

$$\frac{1}{2}\frac{w}{g}v^2 = wh,$$

or

$$v^2 = 2\,gh,$$

or

$$v = \sqrt{2\,gh},$$

where v is the speed of water at the orifice. This is another way of stating Torricelli's theorem. This expression for the speed of the jet can also be derived from Bernoulli's theorem.

154. Transference of hydraulic energy of water to machines. The modern water turbine is designed to extract a maximum amount of the energy in water. So well has this been done that efficiencies as high as 90 per cent have been obtained. This high efficiency has been attained only after a thorough study of the principles of hydraulics and by the aid of experimental investigations carried out in coöperation with the theoretical development. Some of the important things are obvious. For example, (1) the pipes and tubes are made of such shape that the loss of energy due to friction will be small, and (2) the water must leave the discharge tube with as small a speed and pressure as possible. The potential energy is changed into kinetic energy either before entering the wheel, or turbine, or while passing through it. Sometimes the turbine is mounted at a contracted part of the tube carrying the water. At this contraction the kinetic energy is a maximum and the pressure is very low (sect. 146). It should be obvious why turbines are usually located at a place of low elevation, near the level of the stream which carries the water away.

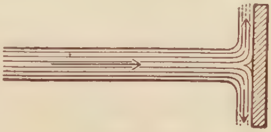


Fig. 116

In the old-fashioned water wheel the water struck a flat blade. This was not an efficient arrangement, as will be shown by the following considerations: When a jet of water strikes the blade shown in Fig. 116, the water spreads out over the surface of the blade and continues moving with the speed of the blade. In that case it does not lose all its kinetic energy. If, however, the jet strikes a curved blade, as shown in Fig. 117, the water will be deflected back from the blade. It will then have a *forward* velocity less than that of the blade. At a certain speed of the blade the water will have very little velocity left. In such a case it loses nearly all its kinetic energy.

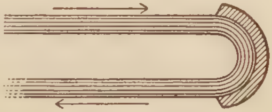


Fig. 117

155. Water wheels and turbines. Water wheels and turbines differ very much in type and design, and the details are too technical to be explained here. When the "head" is small, the speed that can be developed is small, and the turbine must be designed to run at low speeds and to use a large quantity of water. Thus the 10,000-horse-power turbines at Keokuk, Iowa, work under a head of about 32 feet, and run at the low speed of 58 revolutions per minute. But when the head is high, the wheel must be designed for a smaller quantity of water flowing at higher speeds.

In this case "impulse wheels" are sometimes used. In these the energy of the water is converted into kinetic energy of a high-speed jet which strikes moving vanes.

156. Hydraulic ram. A hydraulic ram placed in or near the bed of a stream can raise water to a reservoir at a considerable elevation. Water running through a long pipe flows out through the valve *V* (Fig. 118). This valve is so constructed that it will close when the water flowing through it has attained a certain speed. Just before the valve closes, the water has considerable momentum; the longer the pipe and the greater the speed, the greater is the momentum. When the valve closes, this momentum must disappear. The destruction of this momentum requires an increased pressure in the pipe, for forces are always necessary to overcome momentum. The increased pressure will open the other valve, and a relatively small amount of water will be forced through under this high pressure. As soon as the pressure is relieved, this valve closes; and valve *V*, either on account of its weight or through the action of a weak spring, opens, and a flow begins again. As soon as the water in the pipe again gains sufficient speed, the valve *V* closes, the pressure is raised, and more water flows into the discharge chamber. In this way, at regular intervals, water under high pressure passes into the chamber; and since the pressure there is high, water can flow to a higher level. The principle involved is that an increased pressure is developed when flowing water is suddenly stopped.

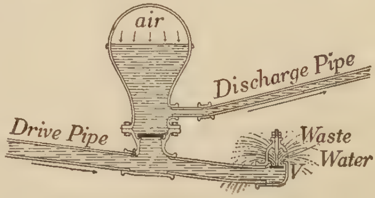


FIG. 118. The hydraulic ram

In an ordinary city-water supply, when a faucet is opened the water in the pipe acquires momentum. If the faucet is quickly closed, there may be a sudden development of high pressure which produces a sharp hammering sound, and which may wreck the plumbing. To prevent this there is usually placed near the faucet a vertical dead-end pipe which is partly full of air. As the pressure rises the air decreases in volume and thus stops the flow in the pipes more gradually.

157. Pitot tube. If a tube shaped as in Fig. 119 is placed in running water with the open end, *A*, pointing against the current, the momentum of the running water will cause a rise of pressure in the tube. This increased pressure may be measured by the displacement of the mercury in the bend. *Pitot* tubes based on this principle are frequently used to measure the speed of running water. A slight modification is sometimes used in airplanes, in order to measure the speed of the airplane with respect to the air.

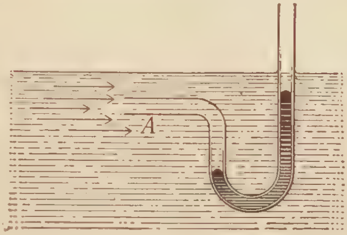


FIG. 119

PROBLEMS

1. Water flows through a horizontal pipe, 6 in. inside diameter, with a velocity of 4 ft./sec. The water from this pipe flows into a smaller pipe, 4 in. inside diameter. What will be the velocity in the 4-inch pipe?

2. What is the pressure in the 4-inch pipe of Problem 1 if the pressure in the 6-inch pipe is 40 lb./sq. in. and if there is no change in elevation?

3. A tube shaped like that of Fig. 105 has water flowing through it. The speed at *A* is 4 ft./sec., and the pressure is 2 lb./sq. in. (in excess of atmospheric pressure). What must be the speed at *B* in order that the pressure will be equal to atmospheric pressure?

4. Water is flowing through a small pipe from a water tank with a speed of 20 ft./sec. If the pressure in the tank on the same level as the outlet is 30 lb./sq. in. (above atmospheric pressure), find the pressure in the small pipe (friction neglected).

5. With what speed would water flow from a small opening in the side of a tank out into the air if the pressure inside the opening were 40 lb./sq. in. (friction neglected)?

6. With what speed would water flow out of a small opening in the side of a tank 16 ft. below the water level (friction neglected)? ($g = 32$ ft./sec.^{2})

7. How much work must be done by a pump on each cubic foot of water that is raised 20 ft. and forced into a main where the pressure is maintained at 30 lb./sq. in.?

8. How much work must be done on each gallon of water if it is raised 30 ft. and forced into a main at a pressure of 50 lb./sq. in.? (7.5 gal. = 1 cu. ft.)

9. Water is drawn from a hydrant at the rate of 1 gal. in 5 sec. How much power is required at the pumping-station to maintain a pressure of 50 lb./sq. in. at the station?

10. If the efficiency of a water motor is 50 per cent, how much water per minute must be supplied to it at a pressure of 30 lb./sq. in. in order that it may develop one quarter of a horse power?

11. State in which case energy is the more expensive: at 10 ¢ per kilowatt-hour or at 10 ¢ per hundred cubic feet of water at a pressure of 40 lb./sq. in.

12. A water turbine has an efficiency of 80 per cent and develops 100 H. P. How many cubic feet of water must be supplied per hour from a reservoir 100 ft. higher than the turbine?

13. How many cubic feet of water must be supplied per hour to the turbine of Problem 12 if the supply has a pressure of 60 lb./sq. in.?

CHAPTER XIV

THE MOLECULAR AND ATOMIC THEORY. SPECIAL PROPERTIES OF SOLIDS AND LIQUIDS DUE TO MOLECULAR ACTION

The molecular and atomic theory of matter, 158. Evaporation, 159. Molecular forces, 160. Surface tension, 161. Measurement of surface tension, 162. Cause of surface tension, 163. Capillarity, 164. Capillary action in the soil, 165. Effects of surface tension on floating bodies, 166. Diffusion of liquids, 167. Diffusion of solids, 168. Diffusion of gases, 169. Diffusion through porous walls, 170. Osmosis; osmotic pressure, 171.

158. The molecular and atomic theory of matter. It was stated in Chapter IV that ordinary bodies are built of molecules. The molecule in turn is built of atoms, the water molecule being a cluster of two hydrogen atoms and one oxygen atom. There are about eighty-seven kinds of atoms which have different chemical properties. Different kinds of molecules are built of different combinations of these atoms. A substance that contains only atoms with the same chemical properties is called an *element*, while one built of molecules that contain atoms with different chemical properties is called a *compound*. Gold, silver, iron, oxygen, and hydrogen are elements; wood, water, common salt, and acids are compounds. In general the structures of molecules are so firm and compact that ordinary mechanical treatment of matter does not break them up.

Long ago the chemist was able to measure the *relative* masses of atoms and molecules. To say that the atomic mass of the nitrogen atom is 14 and that of the oxygen atom is 16 means only that their masses are proportional to these numbers. In recent years tremendous strides have been made in discovering the properties and structure of molecules and atoms. Unexpected and practically conclusive evidence of the correctness of the molecular theory has been furnished by work in the discharge of electricity through rarefied gases and by the study of the properties of radioactive substances and their radiations. Methods have been devised for determining the absolute mass of atoms and molecules. But, more important yet, a body smaller than an atom has been discovered—the *electron.* As a result the theory of the molecule has been extended, and now attempts are being made to explain the structure of the atom. It seems probable that the atom consists of a nucleus with electrons as satellites revolving around it. This nucleus is *very* small,—small compared even with an atom. Evidence seems to indicate that it is about one millionth of the size of an atom. The atoms of different elements differ from one another in the size of the nucleus and the number of satellites (electrons). An atom of greater atomic mass has a nucleus of greater mass and a larger number of satellites. A great deal has been found out about the electron, and it plays so conspicuous a part that the "electron theory" of some of the phenomena of heat, electricity, magnetism, and light has become a very prominent part of the study of those topics.

It was formerly thought that all atoms which had the same chemical properties had the same mass. Now it is thought that chemical properties depend on the number and arrangement of the satellites around the nucleus, and not on the mass. It has, in fact, been discovered that in some cases atoms of different atomic mass have the same chemical properties. For example, ten or a dozen different atoms have the chemical properties of lead. Those atoms which have different atomic masses but the same chemical properties are called *isotopes.* Common lead is a mixture of several kinds of atoms, all having the same chemical properties but different atomic masses. Chlorine is a mixture of two isotopes of atomic masses 35 and 37. Evidence is rapidly accumulating which shows that this is true of many of the so-called elements.

Considerable progress is being made in the study of the structure of the nucleus of the atoms. Present information leads to the view that the nuclei of atoms are built of three different kinds of units: the nucleus of the hydrogen atom, the nucleus of the helium atom, and the electron.

Recent investigations, too abstruse for explanation here, indicate that the following are close to the correct values:

Mass of a hydrogen atom is 1.66×10^{-24} *grams.*

The number of molecules per cubic centimeter in a gas at atmospheric pressure and 0°C. is 2.70×10^{19}.

A method for finding the mass of a hydrogen atom is indicated in Chapter XXXI. The approximate mass of any atom may be found by multiplying its atomic mass by the value given above for the mass of a hydrogen atom. The numbers given above are far beyond the range of our comprehension. For example, the number of molecules in 500 cubic centimeters (about 1 pint) of air at ordinary pressure is $500 \times 2.7 \times 10^{19}$, or 1.4×10^{22}. Suppose that a flask of 500 cubic centimeters could have the air entirely removed from it and then be sealed up. Suppose, further, that a minute hole could now be made in the flask, and that the air should flow in at the rate of 10,000,000, or 10^7, molecules per second. How long would it take to fill the flask to ordinary atmospheric pressure? The student can verify the following:

Since in one year there are approximately 32×10^6 seconds, the number of molecules flowing into the flask in one year would be $32 \times 10^6 \times 10^7$, or 3.2×10^{14}. Hence the number of years required to fill the flask would be

$$\frac{1.4 \times 10^{22}}{3.2 \times 10^{14}} = 0.4 \times 10^8 \text{ (approximately).}$$

But this is equal to forty million years!

159. Evaporation. The evaporation of water is a molecular process. In the interior of water the molecules are in a constant state of motion, some of them breaking through the surface and forming water vapor in the space above. If the molecules in the liquid were at rest and not in a continual state of motion, there would be no reason for supposing that they would leave the liquid. All liquids evaporate, some of them very rapidly (for example, alcohol and gasoline); others, such as heavy oils and mercury, evaporate slowly.

Some solids will evaporate at noticeable rates; for example, gum camphor and iodine. The evaporation of solids is often called *sublimation*. The so-called permanent metals, such as iron, copper, and tin, do not evaporate while in the solid form, but will evaporate when heated into liquids.

After one learns that substances are made of molecules and that these molecules are probably not in contact, it is natural to wonder why all substances do not fall to pieces, or evaporate at very high rates. The molecules must exert attractions on each other.

160. Molecular forces. Molecules are so extremely small that the attraction between them produces no appreciable effect unless they are relatively close together. In order to obtain an attraction which can be detected it is necessary to bring a large number of molecules close together. If two very clean and smooth pieces of lead are pressed tightly together, they will stick. Two pieces of iron may be brought into such close contact by welding that they practically become one piece. Two pieces of lead or iron simply brought together do not show this attraction, for they touch at only a few points. Usually there are dust particles, and even when these are removed there may be left a thin air film. Two pieces of clean plate glass when laid one on the top of the other will have a comparatively thick film of air between them. This is the reason the top one slides so easily over the other. When one strikes the two together, with the surfaces parallel, there is not the metallic sound that should be obtained from glass striking glass, but a distinctly muffled sound due to the air film. This film can be partly removed by pressing the glass plates together and, at the same time, sliding one around on the other. When this film becomes very thin, the plates will cling together. When the experiment is tried in a vacuum, where complications due to atmospheric pressure are to a large extent removed, the attractive force between the plates is shown more conclusively.

That molecular forces are often very great can be shown when the contact between surfaces is sufficiently intimate, as in welding or soldering. It is found that a considerable amount of work must be done to make water evaporate; that is, to tear the molecules apart. (In the chapters on heat it will be shown that energy must be added to water when it evaporates.)

Sometimes a distinction is made between the case where the attraction is between molecules which are alike and the case where it is between different kinds of molecules. The force between like molecules is called *cohesion* and that between unlike molecules is called *adhesion*. Water is held together by cohesive forces, but mucilage sticks to paper by adhesive force. However, this distinction is of little importance.

The following experiment gives an illustration of molecular forces: A glass plate, about 8 or 10 centimeters in diameter, is suspended from one arm of a balance, with its surface horizontal. After the plate is balanced by weights on the other side of the balance, a dish of water is placed under the plate, and the beam is tipped until the glass plate rests on the water. It will be found that additional weights must now be added on the other side of the balance before the glass plate is torn off the water surface. Just before the plate leaves the water it will be noticed that the water under the plate is lifted up by the adhesive force of the water and glass molecules (Fig. 120). When sufficient weight has been added to the other end of the beam to pull the glass off, it will be found that the glass has come away *wet*. The weight which has been added is the force necessary to tear the water molecules apart. The fact that the glass comes away wet shows that the attraction between the glass and the water molecules is appreciably greater than that between the water molecules.

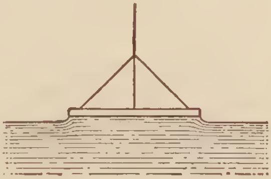


FIG. 120

If the experiment is repeated, using mercury and a clean, dry glass plate, a greater force will be required to pull the glass away from the mercury than from the water. But the glass will now come away dry, showing that the attraction between mercury molecules is greater than that between glass and mercury molecules.

Water will wet some surfaces and will not wet others; for example, that of a wax candle. Mercury will not wet most surfaces. Whenever the cohesion between the molecules of any liquid is greater than the adhesion of the liquid molecules to a surface, then that surface will not be wet by that liquid.

There is an attraction between air molecules and those in the surfaces of bodies exposed to the air. The effect of this is shown when one attempts to produce a high vacuum. Even after a high vacuum is attained an appreciable amount of air clings to the inside walls of the X-ray bulb, or whatever it is that is being emptied of air. Heating the walls of the bulb and the inclosed metal terminals will usually drive off enough air to cause an appreciable increase in the pressure in the bulb.

161. Surface tension. The surface of water or of any other liquid has a tendency to contract. Water sprayed on a dusty surface or mercury spilled on the floor has a tendency to assume a spherical shape. Raindrops, a drop of olive oil suspended in a mixture of alcohol and water, and falling drops of molten metal are all spherical. They assume this shape because the surface of the liquid, in contracting, forces the drop into the shape in which it has the least surface. This must be a sphere, for a sphere is that geometrical figure which, for the same volume, has the least surface. The mist or dewdrops on spider webs or on the downy leaves of plants are minute spheres. They behave as if they were inclosed in elastic bags. The explanation of this tendency for the surfaces of liquids to contract will be given later. The name given to this contractive force is *surface tension.*

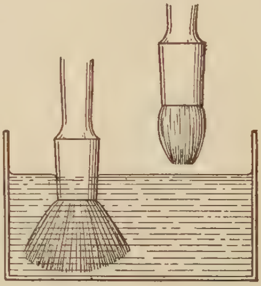


Fig. 121

There are many things we commonly observe that are due to surface tension. A small tumbler of water can be filled until the surface of the liquid is higher than the edges of the vessel. The surface tension of the liquid (that is, the forces which tend to contract the surface) prevents the water from spilling.

While a brush is under water, the hairs do not cling together as they do when it is raised out of the water (Fig. 121). When the brush is out of the water, the surface tension of the water film on the brush draws the hairs together, reducing the surface of the water; but under the water there is no such action. This experiment is easily performed; a shaving-brush and a tumbler of water will do. Wet garments cling to one when he is out of the water but not while he is under the water.

A clean steel needle, when carefully placed on the surface of water, will float. The floating needle does not break through the surface, but pushes it down (Fig. 122) until a relatively large volume of water is displaced. Some insects can move about on the surface of water, being supported by the surface tension. However, if they get wet and get under the surface, the surface tension tends to hold them under, and they may have great difficulty in getting out.

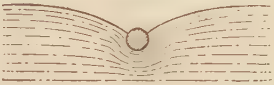


Fig. 122

A clean, dry sieve of rather closely woven wire may hold water. In that case the wire is not wet, and the surface film holds the water. If, however, a moistened finger is touched to the lower side of the wire gauze, water will run through at that place. (To prevent the water from breaking through when pouring it in, a small piece of paper should be placed in the bottom of the sieve.) The same sieve will float on water. More certain results will be obtained if the sieve is first dipped in very hot molten paraffin and the excess of wax quickly shaken off when it is lifted out.

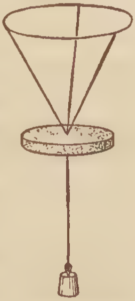


Fig. 123

The wire ring, cork, and lead sinker of Fig. 123 are connected by rigid wires and so adjusted that, when placed upright in water, they will float with the cork almost submerged. When the ring is pushed entirely under the water, the surface tension of the water surface will hold it under; but the surface of the water will be pushed up by the ring. Looking at this and remembering that the surface layer of water behaves as a stretched membrane, one can see why there is a *downward* force acting on the wire ring. But the most striking part of this experiment is yet to come. If, while the ring is under the surface, a few drops of ether fall on the water, the ring will jump up through the surface. The surface tension of a mixture of ether and water is *less* than that of pure water and is not strong enough to hold the ring under.

The relatively weak surface tension of soapy water may be shown by the following experiment: A very fine, light powder, such as lycopodium powder, is sprinkled over the surface of clean water. A wire or piece of glass rod that has been wet with soapy water is touched to the surface of the water. Instantly the powder is swept away from the point touched. The strong surface tension of the pure water pulls the weaker, soapy water out over the surface of the clean water.

The surface tensions of oils are less than that of water; hence oils floating on the surface of water will spread very quickly over the surface. If there is any oil or grease in the water, one will see a thin film of it spread over the surface.

Since the surface tension of hot water is much less than that of cold water, a drop of oil may lie on the surface of hot water (Fig. 124). The tensions of the three surfaces at the place where they meet are shown by the vectors. The drop takes such a shape that the sum of these vectors is zero. These floating drops are seen in hot soups or broth. When, however, the soup gets colder, the tension of the water gets greater, and the oil is pulled out over the surface. A spoon and its contents are then covered with the oil, and the soup has a greasy taste. When some part of a lubricated bearing gets hot, the surface tension of the lubricant is weakened at that place, with the result that the lubricant is pulled away to cooler parts. This may cause a hot-box.

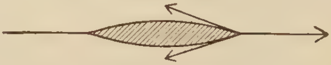


Fig. 124

Small fragments of camphor floating on the surface of clean water move about with very irregular spinning motions, sometimes for hours without stopping. Camphor dissolves only to a small extent in water. But the irregular shape of a particle will tend to make it go into solution faster on one side than on the other. The result is that the surrounding water is more impure at one side than at another, and hence the surface tensions are not equal on all sides of the camphor particle. These unequal forces cause the motion of the particle.

The surface tension of a soap film always tends to cause the film to contract. This is the reason a soap bubble is spherical in shape. The tendency for a film to contract may be shown in several ways. If a soap bubble is blown on the end of a pipe or a glass funnel, it will slowly contract, driving air out through the stem. This current of air is easily shown by its effect on a candle flame. The tension in a soap film may also be shown by the following experiment: A ring of wire is dipped into some soap solution. By carefully lifting it out, a film is left extending over the ring. A piece of wet thread with a loop in the end is placed in this film. If the film inside the loop is punctured, the film on the outside will stretch the loop into the circular form shown in Fig. 125.

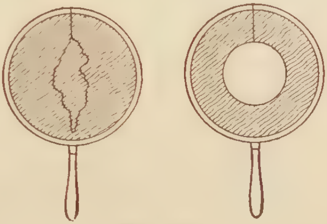


FIG. 125

162. Measurement of surface tension. A very simple form of apparatus serves to measure the surface tension of a liquid. A piece of wire is bent into the shape shown in Fig. 126. This wire is dipped down into the liquid, say water, until it is completely submerged. If it is gently raised, a film (not shown in the figure) will be formed between the horizontal portion of the wire and the surface of the liquid. This film pulls downward along the horizontal portion. If the wire is suspended from a balance, the downward pull can be measured.

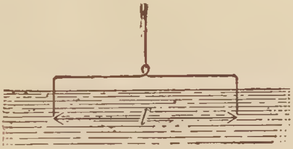


FIG. 126

The tension of a membrane (for example, a drumhead) is, by definition, equal to the pull the membrane exerts on a unit length of the bounding edge of the membrane. Hence the surface tension of a liquid is the pull of one of the surfaces of the film on a unit length of the wire of the experiment given above. If the measured total force is F, then, since there are two surfaces to this film, the formula for the tension S is

$$S = \frac{F}{2l},$$

where l is the length of the horizontal portion of the wire frame.

In the table on the following page, β is the temperature coefficient. The value of the surface tension at any temperature t (centigrade) may be computed by the formula

$$S = S_0 - \beta t.$$

SURFACE TENSIONS

(At 0° Centigrade in Dynes per Centimeter)

	S_0	β
Alcohol	25.3	0.09
Benzine	30.6	0.132
Ether (ethyl)	19.3	0.115
Mercury	527.	0.379
Water	75.6	0.14

163. Cause of surface tension. Surface tension is a force tending to contract the surface. Why does a surface of a liquid tend to get smaller, to have less area? It will be shown that there are forces which are continually trying to pull into the interior of the liquid all the molecules which lie in the surface of the liquid. Any attempt to remove the units which make up the surface produces a tendency to decrease the area of the surface.

A molecule in the interior of a liquid is surrounded on all sides by other molecules, and the mean resultant molecular force on it must be zero, for it is attracted as much in one direction as in the other. In considering the action of the surrounding molecules only those very near need be taken into account, for the range of molecular attraction is small. Hence molecules which are even close to the surface of the liquid are attracted equally in all directions. But this is not true of the molecules in the surface layer. Above a molecule on the surface there are only air molecules, and these are relatively far away, so that the resultant of all the molecular forces acting must be toward the interior of the liquid. Hence there is a tendency to pull the molecules out of the surface, which causes that surface to get just as small as possible, behaving as if it were covered with an elastic film.*

164. Capillarity. Everyone is familiar with the fact that water will rise through the interstices of a towel when one end hangs in water. It is equally well known that oil rises in lamp wicks, and melted wax in the wick of a candle. Another simple case is the rise of water in narrow tubes, usually called capillary tubes. This phenomenon of the rise of liquids in small tubes and in small interstices is called *capillarity*.

---

*According to a modern chemical theory of surface tension the attraction between the molecules is a chemical force (see Lewis, A System of Physical Chemistry, Vol. I, p. 474).

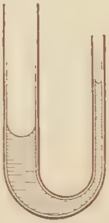


FIG. 127

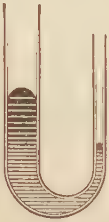


FIG. 128

Fig. 127 shows the relative heights at which water will stand in two connected glass tubes when one is much smaller than the other. Fig. 128 shows the relative heights when the tubes are filled with mercury. In this case the mercury stands at a lower level in the small tube than in the large one. Whether a liquid rises in the smaller tube or not depends on whether the liquid wets the tube. A clean glass tube is wet by water, and water rises in the tube. Mercury does not wet glass, and the surface is depressed. The surface of water in a glass tube which has been coated with wax will be depressed.

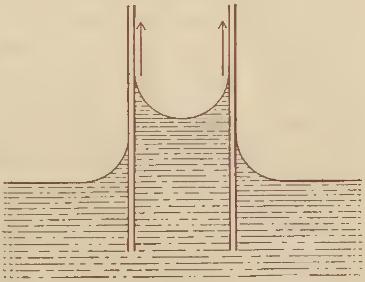


FIG. 129

The *cause* of *capillarity* is not difficult to understand. It depends on the *surface tension* of the liquid. Fig. 129 shows, in exaggerated cross section, a small tube dipping into a liquid which wets it. Such a tube is covered by a thin film of the liquid for some distance above the water. The surface of this film and the surface of the water with which it is in contact are under tension and tend to contract. On the outside of the tube this tension changes the surface into the curved shape shown in the figure, and lifts a small amount of water; on the inside the film lifts the water higher. The vectors in the figure show the direction of the lifting forces. The total upward force on the inside will depend on the length of the edge of the film (that is, on the inside circumference of the tube), while the weight of water lifted is proportional to the area of cross section of the tube. If a smaller tube is used, the cross section is decreased more than the circumference; for the one depends on the square of the diameter, and the other on the first power. Hence we should expect the water to be lifted higher in a smaller tube. This may be proved algebraically as follows:

The surface tension S is the vertical pull on each centimeter of the edge of the film; that is, of the circumference of the tube. If r is the radius of the tube, $2\pi r$ is the circumference. Hence

$$\text{Upward force} = 2\pi r S.$$

The volume of the liquid lifted is approximately $h\pi r^2$, where h is the height the water is lifted. The weight of water lifted is the product of volume, density d, and g.

$$\text{Weight lifted} = h\pi r^2 dg.$$

Equating the upward force to the weight lifted,

$$2\pi r S = h\pi r^2 dg,$$

or

$$S = \frac{hrdg}{2} \tag{81}$$

and

$$h = \frac{2S}{rdg}. \tag{82}$$

Or, the height varies inversely as r, the radius of the tube.

Equation (81) may be used in computing the surface tension S of a liquid when h and the radius r of a tube are measured.

Strictly speaking, the water is not lifted by the surface film. The upward force of the film causes a decrease in the pressure under the surface; above the surface the pressure is that of the atmosphere. Hence in the liquid under the film the pressure will be less than the atmospheric pressure, and the liquid will be forced up by the pressure on the outside of the tube.

The student should now be able to explain why the surface of mercury in a glass tube is depressed. Many of the peculiarities of mercury are due to the fact that its surface tension is very large.

165. Capillary action in the soil. Water is brought up to the surface of the soil partly by capillary action and partly by diffusion of water vapor. If the soil is loosely packed, water will not rise so rapidly by capillary action. Cultivation of the soil in summer is partly to form a dust mulch, thus checking the capillary action.

In the so-called dry farming, moisture in the ground is conserved by intensive cultivation. However, such cultivation may aid the circulation of the water vapor in the soil to such an extent that it may actually dry out faster (see sections 169 and 276). This is a subject the experimental investigation of which is by no means complete.

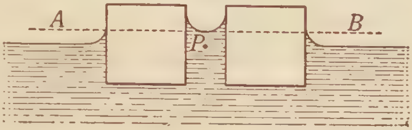


FIG. 130

166. Effects of surface tension on floating bodies. Two floating bodies which are wet by any liquid will attract each other when they come near together. Two pieces of cork floating in water will, after a time, drift together. One can tow a cork in a vessel of water by this attraction between two corks. Sticks, logs, and other drifting material show a tendency to cling together. This can be understood by referring to Fig. 130. Between the two floating objects the water level is higher than outside. At the point P, which is in the general level of the liquid surface, the pressure must be atmospheric; that is, it is at the same pressure as distant points on the same level. At points in the liquid above P the pressure must be less than at P, for the pressure in a liquid decreases with height. But on the outside of these bodies, say at the height of the line AB, the pressure is atmospheric; hence it is larger than that on the inside. The external pressures being greater, there are forces pushing the two floating bodies together. In a somewhat similar way it can be shown that two floating bodies which are not wet by the liquid will be forced together (Fig. 131). But in the case of two bodies one of which is wet and the other not, there will be repulsion (Fig. 132).

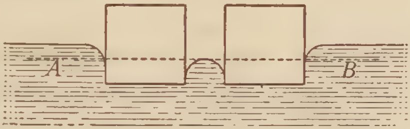


FIG. 131

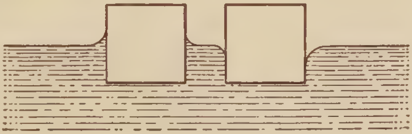


FIG. 132

167. Diffusion of liquids. A cylindrical glass vessel (Fig. 133) is partly filled with water. Then, by means of a glass tube, a solution of copper sulfate is poured into the bottom of the cylinder. If this has been carefully done, there will be a sharp line of separation between the two liquids. In a few days the sharp boundary will have disappeared, and the blue copper-sulfate solution will have moved upward into what was pure water. The copper-sulfate solution is heavier than water; and if it were an inert mass, it would stay at the bottom. But a liquid is not an inert mass: the heavy molecules of copper sulfate and the lighter molecules of water mix on account of their motion. This mixing, produced by molecular activity, is called *diffusion*. If, instead of pure water, the cylinder of Fig. 133 were filled with a solution of blue litmus in water, and a dilute solution of nitric acid poured into the bottom, the upward diffusion of the acid molecules could be seen by the change that the acid produced in the color of the litmus. In this case the diffusion will be much faster than with copper sulfate.

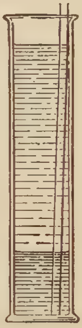


FIG. 133

Diffusion in liquids is always a slow process; but if sufficient time is given, the diffusion will continue until the whole is practically uniform in composition.

168. Diffusion of solids. A very slight diffusion takes place in solids. In an experiment of Roberts-Austen pieces of lead and gold with very clean surfaces were brought into contact. After four years lead was detected at a depth of 5 millimeters in the gold. At a depth of three fourths of a millimeter in the lead there was enough gold to have paid to extract it if done on a large scale. At high temperatures the action is much more rapid. At 500° C. gold diffuses through lead about as fast as common salt (sodium chloride) diffuses through water at 18° C. This is undoubtedly due to the fact that molecules move with greater velocities at higher temperatures.

169. Diffusion of gases. In gases the molecules move with greater freedom, and diffusion takes place at a very high rate. A small quantity of pungent gas released into a still room can be quickly detected in other parts of the room. The different gases which compose the atmosphere differ in weight; yet on account of diffusion they are usually uniformly mixed. Water vapor coming from a water surface diffuses very rapidly through the air. It is often the rate of diffusion that determines the rate of evaporation. In the absence of wind, diffusion prevents the air next to the water from becoming so saturated with water vapor that evaporation is stopped.

For the same temperature and pressure the average speeds of the molecules of different gases vary inversely as the square root of the density of the gas. Light gases will therefore diffuse much faster than heavier ones.

170. Diffusion through porous walls. When hydrogen and air are separated by an unglazed earthenware wall, the hydrogen diffuses through the small interstices faster than air. This is well shown by the following simple experiment: The open end of a cup of porous earthenware is closed by a stopper through which a glass tube passes. The other end of the tube dips (Fig. 134) into a dish of colored water. A glass beaker is inverted over the porous cup. By the aid of a rubber tube, ordinary illuminating gas, which is largely hydrogen, flows from the gas mains up into the beaker. As this gas is lighter than air, it will stay for a short time in the inverted beaker. Just as soon as the gas is turned on, the pressure inside the porous cup rises, causing the air in the glass tube to bubble out through the water. The cause of this is that gas diffuses in through the walls of the porous cup so much faster than air diffuses out that an increase in pressure is produced in the cup. If the beaker is now removed, the gas in the cup will diffuse out more rapidly than air goes in, and the pressure will fall below ordinary atmospheric pressure. This is shown by the rising of the water in the glass tube. Diffusion through membranes can be used for partially separating mixtures of gases. For example, if hydrogen and oxygen are mixed in a vessel which has a porous wall, the hydrogen will diffuse out faster than the oxygen.

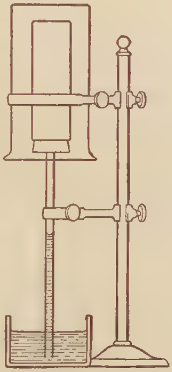


FIG. 134

In the lungs it is diffusion through the surrounding membranes that carries carbon dioxide out and oxygen in.

Metals, when heated, may become so porous that gases will diffuse through them.

Reference has already been made in section 165 to the diffusion of water vapor through the soil. In loose soils it seems probable that the water deeper down evaporates, and that this vapor is carried through the loose upper part out into the air. In some cases this is the chief method by which the ground dries out (see also section 276).

171. Osmosis; osmotic pressure. Liquids also diffuse through porous membranes. In 1748 Abbé Nollet filled a bladder with alcohol and immersed it in water. The water diffused through so much faster than the alcohol that the bladder swelled almost to the bursting-point. The simpler facts of diffusion through membranes can be shown by the following experiment: The large end of a thistle tube is covered with a piece of animal membrane, such as the skin of sausage, and the bulb is filled with a sirup. The bulb is then immersed in water, as shown in Fig. 135. The water molecules will pass through the membrane more rapidly than the molecules of the sugar solution, and in half an hour this sirup will rise in the tube a considerable distance above the level of the water. This process is usually called *osmosis*.

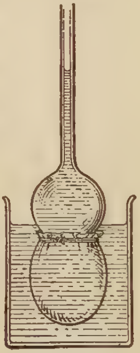


FIG. 135

If the outer vessel also contains a sirup, the osmotic action will always tend to cause water to pass into the more concentrated solution. The tendency will be to make the two solutions of equal concentration.

Theoretically, at least, the solution will rise in the tube until the hydrostatic pressure is sufficient to stop the water from diffusing through the membrane. The pressure which is just sufficient to stop the action is called the *osmotic pressure* of that solution. Instead of actually permitting the sirup to keep on rising in the tube until its hydrostatic pressure stops the action, the usual way is to apply sufficient pressure (for example, by a piston in the tube) to stop the osmotic action. Animal membranes are not strong enough to stand the very large pressures which are often necessary to stop the action. With special types of membranes very great pressures have been developed,—pressures of many atmospheres.

Osmosis plays a very important part in plant and animal growth. The roots of a plant contain sap, a more concentrated solution than the soil water. The surface of the root acts as the membrane, and water passes into the root. Osmotic action takes place through membranes in the bodies of human beings and animals and forms a very important part in growth and other activities.

PROBLEMS

1. Why does a damp cloth absorb water more quickly than a dry one?

2. How high will water rise in a tube which is 2 mm. in diameter?

3. How much will the surface of mercury be depressed in a tube 4 mm. in inside diameter?

4. Mercury is raised by gas pressure 23.2 cm. in a tube 4 mm. in inside diameter. Compute the error produced by capillary action.

5. (*a*) If the length of the horizontal portion of the wire frame of Fig. 126 is 8.0 cm., with what force will a water film pull downward on it? (*b*) What will be the force if alcohol is used?

6. If the inside diameter of one tube of Fig. 127 is 1.0 cm. and that of the other is 3 mm., what will be the difference in levels (*a*) if the tubes contain water? (*b*) if the tubes contain mercury?

CHAPTER XV

ELASTICITY

Elasticity, 172. Hooke's law, 173. Limit of elasticity, 174. Stress and strain, 175. Young's modulus, 176. Practical use of Young's modulus, 177. Volume elasticity, 178. Modulus of volume elasticity, or bulk modulus, 179. Elasticity of shape, or rigidity modulus, 180. Other cases of elasticity, 181. Ultimate stress or strength, 182.

172. Elasticity. It is often important for one to understand how to compute the amount a wire or rod will stretch, or a beam will bend, under a given load. Not only in the design of large structures, such as bridges or the steel frame of a building, but often in very simple cases it is necessary to compute the extension, bending, or twisting under the maximum load. In this chapter is given an elementary introduction to the fundamental definitions, facts, and laws of the subject of *elasticity.*

173. Hooke's law. If a vertical wire or small rod is clamped at the upper end, and weights are applied at the bottom, results something like the following are obtained:

BRASS WIRE: LENGTH 195 CENTIMETERS, DIAMETER 0.047 CENTIMETERS

Load in Gram-weights	Elongation in Centimeters	Change per 200 Gram-weights
200	0.026	+ .026
400	0.055	+ .029
600	0.087	+ .032
800	0.113	+ .026
1000	0.139	+ .026
1200	0.165	+ .026
1000	0.139	− .026
800	0.113	− .026
600	0.089	− .024
400	0.063	− .026
200	0.034	− .029

An inspection of the table given above, which is taken from actual observations, shows that the elongation is closely proportional to the stretching force. This is a case of a general law due to Robert Hooke. In 1676 Hooke published the law as *Ut tensio sic vis*, which, translated freely, means "The extension is proportional to the force." It should be clearly understood that this law is a summary of the results of experiment. In many cases it is fairly accurate; in many other cases only roughly true, but near enough for most practical purposes.

Laws similar to this are approximately true for cases other than stretching; for example, the compression of solids and liquids, the bending of beams and plates, and the twisting of wires and rods. In some treatises Hooke's law is generalized and is said to apply to all these cases, but originally it applied only to stretching.

In the chapter entitled "Vibratory Motion" it was stated that in any case where the restoring force was proportional to the displacement the vibrations were of the *simple harmonic* type. Another way of stating the same thing is: In those vibrations where the displacements obey Hooke's law the motion is simple harmonic motion. This applies to cases where the "damping," or frictional force tending to stop motion, is small.

174. Limit of elasticity. The experimental data given in the last section show that when the load is removed the wire resumes (approximately) its original length. This is usually true if one does not deform the body too far. If a rod or wire is stretched too much, it will not return but takes a "set." The *range of perfect elasticity* is the range within which the rod or wire resumes its original length when the load is removed. The limit of this range is called the *elastic limit*. In some materials (for example, lead) the elastic limit is very small; in other cases (for example, hard-drawn brass and tempered steel) the range is relatively large. A spring balance is useful because it has a wide range of perfect elasticity.

175. Stress and strain. Consider a vertical wire or a rod which is stretched by a force applied at one end. This force is transmitted over the entire length of the wire, each part being stretched. When one portion of the wire communicates the force to the next, the force is distributed over the cross section of the wire. The effective force is *the force per unit area of the cross section.* This is called the *stress.* In the case of stretching,

$$\text{Stress} = \frac{\text{applied force}}{\text{area of cross section}}.$$

As an example take the case of a wire the cross section of which is .01 square inches, stretched by a load of 25 pounds. The *stress* is 2500 pounds per square inch.

Strain is defined as the stretch per unit length of the wire or rod; or

$$\text{Strain} = \frac{\text{elongation}}{\text{length of rod}}.$$

Strain may be defined as the fractional change of the length. It is numerically equal to the elongation of unit length of the rod. When the elongation and the length are measured in the same units, the strain has the same numerical value whatever units are used.

176. Young's modulus. Hooke's law may now be stated in the following form: For any one material within the elastic limit *strain is proportional to stress.* Hence the ratio of stress and strain is a constant. This constant, not only in this case but in other kinds of elasticity, is called a *modulus.*

$$\text{Modulus of stretch} = \frac{\text{stress}}{\text{strain}}.$$

This modulus is known as *Young's modulus*, and is usually denoted by the letter M. If the elongation is denoted by e, the length by l, the stretching force by F, and the cross section by A, then

$$\text{Stress} = \frac{F}{A}$$

and

$$\text{Strain} = \frac{e}{l}.$$

$$\text{Young's modulus} = \frac{\text{stress}}{\text{strain}} = \frac{\frac{F}{A}}{\frac{e}{l}} = \frac{Fl}{eA},$$

or

$$M = \frac{Fl}{eA}. \qquad (83)$$

The numerical values of Young's modulus are often quite different for different specimens of the same substance, depending on annealing and the age, size, and shape of the specimen. The following table gives approximate values for common materials.

VALUES OF YOUNG'S MODULUS

MATERIAL	DYNES/CM.^{2}	LB./IN.^{2}
Brass (cast)	$7\text{–}9 \times 10^{11}$	$10\text{–}14 \times 10^6$
Brass wire	10×10^{11}	14×10^6
Copper (drawn)	12×10^{11}	18×10^6
Copper (annealed)	10×10^{11}	14×10^6
Iron (cast)	$10\text{–}13 \times 10^{11}$	$14\text{–}19 \times 10^6$
Iron (wrought)	$19\text{–}20 \times 10^{11}$	$27\text{–}29 \times 10^6$
Steel	$19\text{–}21 \times 10^{11}$	$27\text{–}30 \times 10^6$
Glass	$5\text{–}6 \times 10^{11}$	$7\text{–}9 \times 10^6$

177. Practical use of Young's modulus. Suppose that one wishes to know how much load a vertical steel rod one tenth of an inch in diameter and 10 feet long will support with an elongation of only one twentieth of an inch.*

From equation (83) $F = M\dfrac{eA}{l}$,

where $M = 28 \times 10^6$ lb./in.^{2},

$l = 10 \times 12 = 120$ in.,

$e = \frac{1}{20} = .05$ in.,

and $A = \pi r^2 = .00785$ sq. in.

Substituting these numerical values,

$$F = 28 \times 10^6 \frac{.05 \times .00785}{120} = 91.7 \text{ lb.}$$

How much will a load of 10 kilos stretch a brass wire 2 mm. in diameter and 1 m. long?

---

* In this and the following problem the weight of the rod or wire may be neglected.

From equation (83) $e=\frac{Fl}{MA}$,

where $F=(10{,}000\times 980)$ dynes,

$l=100$ cm.,

$M=10\times 10^{11}$ dynes/cm.^{2},

and $A=\pi r^2=.0314$ sq. cm.

Substituting these numerical values,

$$e=\frac{9800000\times 100}{(10\times 10^{11})(.0314)}=.03 \text{ cm.}$$

If one wishes to find how large a rod or wire to use so that the stretch may not be greater than a certain amount, it can be determined from equation (83), provided the load, the length of rod or wire, and the modulus are known. First, equation (83) is solved for A, the cross section, giving

$$A=\frac{Fl}{Me}.$$

Then the known values of the quantities on the right are substituted and the cross section A computed.

If a force applied at the end of a rod tends to compress it, the amount shortened is very nearly equal to the amount the rod would be stretched if the same force were used to stretch it, provided the changes are within the elastic limit. Hence the methods just explained can be applied to the longitudinal compression of rods and bars.

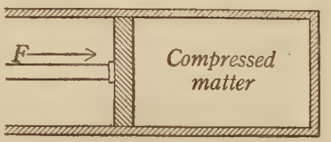


FIG. 136

178. Volume elasticity. Imagine a liquid compressed by some method similar to the one represented in Fig. 136. The compressed substance tends to expand, and reacts with a force on the piston equal to F, the applied force. Such a case comes under the head of *volume elasticity*. The *stress* is the force F divided by the area of the piston. But this ratio is equal to what we call pressure. Hence, in volume elasticity, *stress is numerically equal to pressure.*

The *strain* is defined in terms similar to those of section 175, by saying that it is the fractional change in volume; or,

$$\text{Strain} = \frac{\text{change in volume}}{\text{volume}}.$$

In the cases classified under volume elasticity the forces tending to change the volume act on all sides of the substance. In the case illustrated in Fig. 136 there is not only a force supplied by the piston but also forces from the walls of the container; the pressure is uniform over the entire surface. But if the force of compression is applied in only one direction (for example, in the case of a load acting downward on a supporting pillar), it is an instance of elasticity of stretch, the compression being a sort of negative stretch.

179. Modulus of volume elasticity, or bulk modulus. For solids and liquids it is found that within certain limits the ratio of *volume stress* to *volume strain* is a constant. This constant is called the *bulk modulus*, or *modulus of volume elasticity.*

$$\begin{aligned}\text{Bulk modulus } (K) &= \frac{\text{stress}}{\text{strain}}\\ &= \frac{\text{pressure}}{\dfrac{\text{change in volume}}{\text{volume}}}\\ &= \frac{pv}{\text{change in volume}}, \qquad (84)\end{aligned}$$

where p is the pressure and v the volume.

In the case of gases the ratio of volume stress to volume strain is not a constant. If the gas is maintained at a constant temperature, it can be proved that the bulk modulus, or ratio of stress to strain, is numerically equal to the pressure of the gas, and hence increases as the pressure increases.

Volume elasticity is the only kind of elasticity that liquids and gases have.

In using the bulk modulus it makes no difference what unit is used for measuring volume if the same unit is used for measuring the change in volume. But the value of the modulus will depend on the unit used in measuring pressure. The pressure units commonly used are *dynes per square centimeter*, or *bars, pounds per square inch*, and *atmospheres*. When atmospheres are used, the bulk modulus is numerically equal to the reciprocal of the *compressibility* used in section 34.

BULK MODULI (K)

	Bars	Lb./In.^{2}	Atmospheres
Aluminum	6.7×10^{11}	10×10^{6}	6.5×10^{5}
Copper	13×10^{11}	19×10^{6}	13×10^{5}
Glass	$3\text{–}6 \times 10^{11}$	$4\text{–}9 \times 10^{6}$	$3\text{–}6 \times 10^{5}$
Mercury	2.7×10^{11}	4.0×10^{6}	2.6×10^{5}
Steel	$15\text{–}19 \times 10^{11}$	$22\text{–}27 \times 10^{6}$	$15\text{–}18 \times 10^{5}$
Water	0.23×10^{11}	0.33×10^{6}	0.22×10^{5}

180. Elasticity of shape, or rigidity modulus. Solids differ from liquids in that they tend to preserve their shape. If a rod is twisted, the rod exerts a force opposing the action of twisting. This tendency to preserve the shape is called the *rigidity*. The *modulus of rigidity* is defined in a manner somewhat similar to that used in sections 176 and 179. The details and precise definitions are rather technical and will not be given here.

It has been theoretically proved and experimentally verified that when a rod of length l and radius r is twisted by a torque L applied at the free end of the rod, the angle through which the end is twisted is

$$\phi = \frac{2\,lL}{n\pi r^4}, \tag{85}$$

where n is the modulus of rigidity of the material of which the rod is made. When the modulus can be found in a table, the amount of twist given to a shaft in any machine can be computed, provided the torque and the dimensions of the shaft are known. Such computations are sometimes very necessary.

From equation (85) the simple laws of twist can be deduced: (1) the angle of twist is proportional to the torque; (2) the angle of twist is proportional to the length of the rod; (3) the angle of twist is inversely proportional to the fourth power of the radius; (4) the angle of twist is inversely proportional to the modulus of rigidity.

MODULI OF RIGIDITY (n)

	DYNES/CM.^{2}	LB./IN.^{2}
Aluminum	$2.38–3.36 \times 10^{11}$	$3.4–4.8 \times 10^6$
Brass wire	$3.1–4.0 \times 10^{11}$	$4.6–5.8 \times 10^6$
Copper (drawn)	4.0×10^{11}	6×10^6
Glass	$2.0–3.0 \times 10^{11}$	$3–4 \times 10^6$
Steel	$7.7–9.8 \times 10^{11}$	$11–14 \times 10^6$

181. Other cases of elasticity. All cases of deformation can be stated in terms of the moduli which have been defined. In advanced treatises on elasticity, formulas are worked out for the bending of beams under various conditions, the stability of loaded pillars, the elongation and bending of various types of springs, and many other practical problems. In the case of a beam the bending will depend on the shape of the cross section of the beam and on how the ends are fastened. If the loaded beam is supported at the ends, with the ends free (that is, not clamped), the upper surface of the beam will be compressed and the lower surface stretched. The theory will involve only the elasticity of stretch. When the beam is rectangular in cross section, as in the case of a heavy plank, and a load w is applied at the center, theory gives, as the depression of the center,

$$\text{Depression} = \frac{wl^3}{4\,Mbd^3},$$

where l is the distance between the supports, b the breadth, d the vertical thickness, and M Young's modulus for the material used. If two planks are laid one on top of the other, and securely nailed together, so that one cannot slip over the other, the vertical thickness of the resulting beam is twice as great as that of one plank. But the formula shows that the bending varies inversely as the *cube* of the thickness. Hence the bending would be one eighth of that of a single plank. The beam is much stronger than if it had not been nailed together. For if the planks are not nailed together, we have two beams instead of one, each taking approximately half the load, and the two bending approximately half as much as one plank alone.

182. Ultimate stress or strength. By the term *ultimate stress*, or *ultimate strength*, is meant the greatest stress which can be produced in a body without rupturing it. *Ultimate tension* is the maximum stress that can be applied to stretch a body without rupturing it. It is often called *tensile strength*. The maximum compressional stress that can be applied to a body without breaking it is usually called *compressional strength*. It is customary in specifying the values of all these qualities to state them in terms of force per unit area.

PROBLEMS

1. A load of 3000 dynes is hung by a wire 1 mm. in radius. Compute the stress.

2. A load of 2000 lb. is supported by a rod half an inch in diameter. Compute the stress.

3. A rod 20 ft. long is stretched one fourth of an inch. Compute the strain.

4. A wire 0.5 mm. in diameter and 3 m. long is stretched 3.0 mm. What is the strain?

5. How much will a steel wire 1 mm. in radius and 2 m. long be stretched when it supports a mass of 10 kg.? ($M = 21 \times 10^{11}$ dynes/cm.^{2})

6. How much will an iron rod with radius 0.1 in. and length 10 ft. be stretched by a load of 200 lb.? ($M = 27 \times 10^6$ lb./in.^{2})

7. Compute Young's modulus from the data given in section 173.

8. A rod 100 in. long and 2 sq. in. in cross section supports a stretching load of 30 tons. If Young's modulus is 30×10^6 lb./in.^{2}, compute the elongation.

9. A vertical rod 2 m. long and 0.2 sq. cm. in cross section is stretched 0.08 cm. The load is the weight of 100 kg. Find (*a*) the stress, (*b*) the strain, and (*c*) Young's modulus.

10. What is the fractional change in the volume of a glass ball produced by a pressure of 3000 lb./in.^{2}? ($K = 6 \times 10^6$ lb./in.^{2})

11. A steel shaft 30 ft. long, diameter 1 in., rotates at 300 R.P.M. One end of the shaft is twisted 0.1 radians. (*a*) How much is the torque? ($n = 12 \times 10^6$ lb./in.^{2}) (*b*) What power is transmitted?

PART II. HEAT

CHAPTER XVI

TEMPERATURE AND ITS MEASUREMENT

Temperature, 183. Temperature scales, 184. Thermometers, 185. Maximum and minimum thermometers, 186. The clinical thermometer, 187. Resistance thermometers, 188. Gas thermometers, 189. Thermoelectric thermometers, 190. Measurement of high temperatures, 191. Standard temperatures, 192.

183. Temperature. Our sense organs have taught each of us the meaning of the words *cold* and *hot*. The word *temperature* refers to the same thing as *hotness*, and is used when we have in mind some method for its determination more accurate than our sense organs.

In temperature measurements some physical property of a substance which varies with temperature must be used. In our common thermometers the change in volume of a liquid is used. Some other methods will be described later.

184. Temperature scales. In any temperature scale it is necessary to define at least two *fixed points* on the scale. The two commonly used are the temperature at which pure ice melts and the temperature at which pure water boils under standard barometric pressure. A hundred years ago there were a large number of different temperature scales in use, but this number has diminished until there are at present only three scales in common use. These are the *Fahrenheit*, used chiefly in English-speaking countries; the *centigrade*, used universally in scientific work, and as the domestic scale in some countries of Europe; and the *Réaumur*, used commonly in certain European countries (for example, Sweden, Germany, and Russia) and in some industries elsewhere.

The "fixed points" of these scales are defined as follows:

	CENTIGRADE	FAHRENHEIT	RÉAUMUR
Melting point of ice	0°	32°	0°
Boiling point of water . . . . . at a pressure of 76 cm. . . .	100°	212°	80°

From these values one can see that the interval between the two fixed points is divided in the centigrade scale into 100 divisions, in the Fahrenheit into 180, and in the Réaumur into 80. Hence 100 divisions centigrade = 180 Fahrenheit = 80 Réaumur. These numbers are in the ratio of 5, 9, 4. From this the student should be able to deduce the following rules:

$$F = \tfrac{9}{5} C + 32,$$
$$C = \tfrac{5}{9}(F - 32),$$
$$R = \tfrac{4}{9}(F - 32),$$

where F is the temperature on the Fahrenheit scale, C the reading of the same temperature on the centigrade scale, and R that on the Réaumur scale.

The *absolute*, or *Kelvin* scale, differs from the centigrade only in the position of the zero point, its divisions being the same size. Its zero point is $-273°$ C. The *gas* scale is practically identical with the Kelvin scale. The relation between the centigrade and Kelvin scales is very simple:

$$K = C + 273.$$

The Kelvin scale is used only in scientific work, especially in dealing with very high or very low temperatures and in many computations involving the laws of gases and the laws of radiation.

185. Thermometers. Practically all substances expand with an increase of temperature and contract when the temperature falls. All common thermometers utilize this property. The mercury or alcohol in a glass thermometer rises when the temperature increases because its volume increases more than the volume of the glass container. If glass expanded more than the liquid, the thermometer would fall with a rise of temperature.

The mercury thermometer is more accurate than the spirit thermometer, but its visibility is not so good. The mercury thermometer must have a tube of smaller bore than the spirit thermometer; hence it is usually harder to see. The reason the mercury column must be more narrow is that mercury expands at about one sixth the rate at which alcohol does.

In any careful work the thermometer used should be carefully compared with a standard thermometer. Often in the laboratory the student is expected to test the ice point by submerging the thermometer in crushed ice and to test the boiling point by placing it in the steam arising from boiling water. Practically all the larger laboratories of physical science have standard thermometers which have been carefully tested at the Bureau of Standards or some other testing laboratory.

186. Maximum and minimum thermometers. There are many occasions when it is important to know the highest or lowest temperature that has been reached during a certain period. For example, one might wish to know the lowest temperature that has been reached during the night. A *maximum* thermometer is one that indicates the highest, or maximum, temperature to which it has been subjected since it was set; a *minimum* thermometer is one that indicates the lowest, or minimum, temperature. Fig. 137 shows diagrammatically one form of thermometer which serves as both a maximum and a minimum thermometer. The large bulb *B* and the tube as far as the point *A* contain alcohol or some other suitable liquid. The tube from *A* to *C* contains mercury. Above *C* there is more alcohol, which extends up into *D*, but only partly fills it, thus giving room for expansion. When the temperature rises,

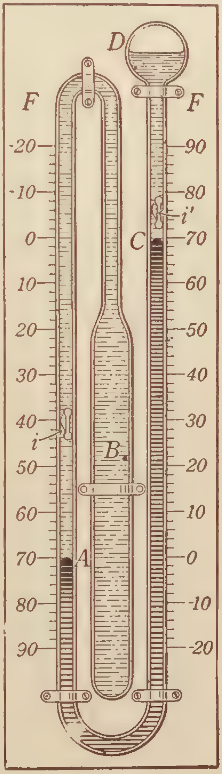


FIG. 137

the liquid in B, expanding, lowers the mercury level at A and raises that at C. There is on each side a scale, but the one at A is numbered downward. The temperature at any time can be read on either scale. Above the mercury, at both A and C, there are placed on the inside of the tube small iron indexes i and i'. Each index is so adjusted that it can slide up and down the inside of the tube, but with enough friction so that its weight will not move it. When the mercury at C rises, it pushes the index i' up the tube; but when the temperature falls, and the mercury at C also falls, the index will be left standing in the tube at the highest point reached by the mercury. Hence this index will give the highest, or maximum, temperature. On the other side the mercury will also push the index up, but only while the temperature is falling. The index on the left side will therefore give the lowest, or minimum, temperature. The alcohol, or whatever liquid is used, wets these indexes and flows past them, exerting inappreciable forces on them. As is well known, mercury does not wet iron, and because of its very large surface tension it can exert considerable force on the indexes. To set the instrument the indexes are jarred down, or pulled down by a small magnet.

187. The clinical thermometer. The clinical thermometer is a maximum thermometer designed to indicate the highest temperature to which it is exposed. Just above the bulb is a narrow constriction through which the mercury flows with difficulty. When the mercury in the bulb is expanding, mercury is forced through; but when the temperature falls, the mercury does not flow back through this constriction, but stands in the tube just as it was at the highest temperature. To set the thermometer the mercury is jarred back through the constriction to the bulb by one or two vigorous shakes.

188. Resistance thermometers. The electrical *resistance* of a coil of wire changes when the temperature of the wire changes. The apparatus needed to measure the electrical resistance of a wire is portable, and the process is not difficult. The change in electrical resistance thus affords a relatively simple method of measuring changes in temperature. It is especially useful for the measurement of temperature in inaccessible places; for example, a coil of wire has been buried in a silo, with the connecting wires extending to the outside. With suitable apparatus the temperature of the fermenting silage is obtained whenever desired.

Resistance thermometers can be used through a wide range of temperature—from very low temperatures up to about 1000° C., or 1800° F. (bright-red heat). For both high and low temperatures the best thermometer coils are made of platinum wire.

189. Gas thermometers. When air or any other gas is heated, it tends to expand; but if prevented from expanding, its pressure will increase. Either the increase in pressure or the increase in the volume of a gas can be used for indicating changes in temperature. On account of the relatively simple laws governing the behavior of gases, gas thermometers are more accurate than any other type. They are not so easy to use, however; hence it is now a common practice, where a high degree of accuracy is desired, to calibrate some other type of thermometer by the more accurate gas thermometer and then to use the calibrated one.

Through a great range of temperatures, from the lowest that are measured up to nearly 1500° C. (white heat), the gas thermometer is the standard thermometer. For very low temperatures helium gas at reduced pressure is used; for high temperatures nitrogen is used. Current practice favors that type of gas thermometer in which the volume of the gas is kept constant and the temperature changes are measured by changes in pressure of the gas.

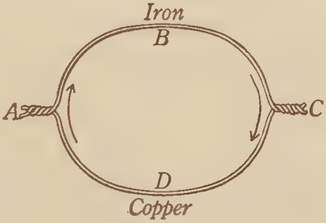


Fig. 138

190. Thermoelectric thermometers. In Fig. 138, *ABC* represents an iron wire and *ADC* a copper one. If the junction *A* of the two wires is heated, and the other junction, *C*, is kept cold, an electric current will flow around the circuit in the direction indicated by the arrows. This is an example of a general rule: *Whenever two wires of different materials are connected together at both ends, an electric current will be developed if the two junctions are not at the same temperature.* When the necessary electrical measurements (which need not be described here) can be made, this effect can be applied to the measurement of the difference of temperature of the two junctions.

The magnitude of the current and its direction of flow depend on the materials used. For ordinary temperatures copper and an alloy called *constantan* (trade name, "advance") are perhaps the best. For high temperatures wires that will not oxidize or readily melt are used. A thermoelectric couple composed of platinum and a platinum-rhodium alloy can be used up to nearly 1600° C.

191. Measurement of high temperatures. In many industries certain processes must be carried on at high temperatures, and it is often an important matter to know when the temperature is correct, for a change of 20° C. may alter the product. Hence a great deal of attention has been paid to the measurement of high temperatures. As has already been stated, the resistance thermometer can be used up to a bright-red heat, and the thermoelectric couple still higher. But the most convenient *pyrometers* (as high-temperature thermometers are called) compare the light or heat radiated from the hot body with that radiated by some other substance. For example, in one type the electric current through an incandescent lamp is changed until the filament of the lamp is of the same brightness as that of the hot body in the furnace, where the temperature is desired. When the electric current through the lamp has been measured, the corresponding temperature can be found from a chart.

192. Standard temperatures. The two fixed temperatures, the melting point of ice and the boiling point of water, are convenient for ordinary ranges of temperatures; but for both low-temperature and high-temperature measurements it is convenient to have a number of known temperatures to use in calibrating, or testing, the instrument one is using. Such low temperatures as the boiling points of liquid hydrogen and oxygen have been carefully determined for this purpose. The melting of many pure metals occurs at very definite temperatures, and these are known. The Bureau of Standards at Washington is prepared to furnish samples of chemically pure metals the melting points of which are as follows:

Tin	231.88° C.
Zinc	419.44° C.
Aluminum	658.68° C.
Copper	1083.00° C.

Other temperatures at which natural processes, such as melting or boiling, take place are accurately known. These temperatures which are known to a high degree of precision are called *standard temperatures.* It is now relatively simple to calibrate any type of thermometer through any desired range by the aid of these standard temperatures.

PROBLEMS

1. Reduce 60° F. to C.; 70° F. to C.; 20° F. to C.

2. Reduce 30° C. to F.; 20° C. to F.; 4° C. to F.

3. Reduce −10° C. to F.; −40° C. to F.; −273° C. to F.

4. Reduce 0° F. to C.; −40° F. to C.; 50° F. to C.

5. Reduce 20° C. to K.; 68° F. to K.; 14° F. to K.

CHAPTER XVII

EXPANSION

Expansion and contraction produced by changes in temperature, 193. Differential expansion, 194. Law of linear expansion, 195. Special cases of expansion, 196. Compensated pendulums, 197. Compensated balance wheels, 198. Volume expansion, 199. Expansion of liquids; relative expansion, 200. Measurement of absolute expansion of a liquid, 201. Expansion of water; maximum density, 202. Forces involved in expansion, 203. Expansion of gases, 204. Expansion at constant pressure, 205. Changes in pressure at constant volume, 206. Summary of gas laws, 207. General gas law, 208. Examples, 209. Limit of application of gas laws, 210. The gas thermometer, 211. Summary, 212.

193. Expansion and contraction produced by changes in temperature. It is not difficult to show that the length of a metal rod increases when the temperature is raised. A simple method can be understood by the aid of Fig. 139. One end of a metal rod is firmly clamped, and the other is allowed to rest on a small sewing-needle which is lying on a smooth surface, such as glass. When the rod is heated, it will expand, causing the needle to roll. A light index fastened to the needle will turn as the needle rolls, thus making the motion visible. If the diameter of the needle and the angle through which it turns arė both measured, it is not difficult to measure the amount of the elongation.

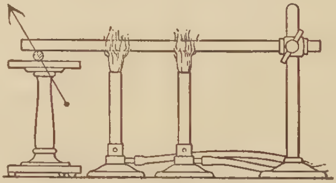


FIG. 139

Another simple method is to send an electric current through a stretched iron wire two or three meters long. If the current is large enough to heat the wire quite hot, one can plainly see the sag of the wire.

A rise of temperature produces expansion, but a lowering of temperature causes contraction. This is plainly seen in either one of the preceding cases when the rod or wire is cooling. The cracks that appear in ice on ponds in severe winter weather are caused by the contraction of the ice in cooling down to a temperature much lower than the one at which it froze.

Expansion or contraction caused by changes in temperature plays an important part in most structures. For example, all long steel bridges must be so built that the structure can expand and contract. In a steel bridge 1000 feet long the change in the length from winter to summer temperatures (that is, from 0° F. to 90° F.) is about one half of a foot. The steel rails on railroads are usually placed with small gaps at the ends to allow for changes in length. Telegraph and telephone lines sag more in the summer than in winter. In laying concrete sidewalks and roadways there must be expansion joints left.

A very interesting phenomenon due to expansion and contraction is the gradual creeping downhill of certain materials. Tyndall describes a lead roof, used on a part of the Bristol cathedral, which, in spite of all efforts to fasten it, slowly slid down, moving eighteen inches in two years. In the daytime the rise of temperature caused it to expand. But it was much easier for it to expand downhill than uphill. At night, when it contracted, it pulled the higher part down, rather than the lower part up; hence it slowly pushed its way downhill. The motion of glaciers on hillsides has to a large extent been explained in a similar way. Lead telephone cables show a similar tendency to move downhill.

194. Differential expansion. If strips of brass and iron are riveted together, the two will not expand and contract equally, for brass expands about 50 per cent more than iron for the same temperature change. The result is that when the riveted strips are heated they will bend as shown in Fig. 140. The essential part of one type of thermostat is two strips of different material riveted together. When heated it bends one way, closing an electric circuit which may close the drafts of the furnace. As the temperature falls, it bends the other way, finally closing a different electric circuit which opens the drafts. It attempts to do what its name indicates—maintain a constant temperature.

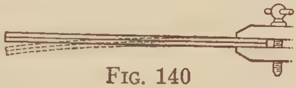


Fig. 140

The breaking of glassware when boiling water is poured into it is caused by differential expansion. If both the inside and the outside of the vessel are heated at the same time, and not too rapidly, the entire vessel will expand at the same time. Thin-walled glass beakers and flasks, such as those used in laboratories, are not so easily broken by sudden temperature changes. Why?

195. Laws of linear expansion. In many cases it is necessary to know beforehand the amount of expansion produced by temperature change. In this section the method of computing the expansion will be explained. Two simple laws can be stated:

1. If the change of temperature is not too large, it is found that *the change in length is proportional to the change in temperature.*
2. *The change in length is proportional to the original length of the body.*

These two facts can be stated by the following algebraic expression:

$$\text{Change of length} = \alpha l_0 t, \tag{1}$$

where l_0 is the original length and t the change in temperature. The constant α is called the *coefficient of linear expansion.* It may be defined as follows: *The coefficient of linear expansion of any substance is the change in length of each unit of length when the temperature is changed 1 degree.* The coefficients of expansion of most substances have been carefully determined by experiment.

The student should see that equation (1) follows from this definition and from 1 and 2 given above. That is, if α is the change produced in unit length for a change of 1 degree, αl_0 will be the change in a length l_0 when the temperature changes 1 degree, and $\alpha l_0 t$ will be the change when the temperature changes t degrees.

If the final length is denoted by l_t, the change in length is

$$l_t - l_0 = \alpha l_0 t. \tag{2}$$

Hence

$$l_t = l_0 + \alpha l_0 t.$$

This may be written in a more convenient form:

$$l_t = l_0(1 + \alpha t). \tag{3}$$

Equation (3) is the formula for computing the length when the change of temperature, the original length, and the coefficient α are known. The following table gives the values of the coefficients of linear expansion for some of the more common substances. In using this table centigrade temperatures must be taken. In all cases given in this table the coefficient is *positive*; that is, the substance expands with rise of temperature. Substances which contract with increase of temperature are very exceptional.

COEFFICIENTS OF LINEAR EXPANSION FOR DEGREES CENTIGRADE

Aluminum	0.000022	Lead	0.000028
Brass	0.000018	Oak (parallel to grain)	0.000005
Copper	0.000016	Oak (perpendicular to grain)	0.00005
Glass (soft soda)	0.000007	Platinum	0.0000090
Glass (lead)	0.0000092	Porcelain	0.000004
Glass ("pyrex")	0.0000032	Quartz (fused)	0.0000004
Ice	0.000051	Silver	0.000019
Invar	0.0000007	Tin	0.000021
Iron (cast)	0.000010	Tungsten	0.0000044
Iron (wrought)	0.000012	Zinc	0.000026

196. Special cases of expansion. *Invar* metal, a nickel-steel alloy, has an extremely small coefficient. For this reason this metal is valuable for clock pendulums, standard meters, and surveyors' tapes. Glass having a very small coefficient of expansion is now made commercially. Hence we now can have ovens with glass windows, glass baking-ware, and, in the laboratory, glass utensils which cannot be easily broken by sudden changes in temperature.

The expansion coefficients of platinum and ordinary glass are nearly the same; hence it is relatively easy to seal platinum wire in glass. When copper is used, unless special processes are employed, the wire breaks loose when it cools. Why? For a long time platinum wires were the only kind that could be sealed into the glass of the ordinary incandescent lamp. Fortunately an alloy has been discovered which has approximately the same expansion coefficient as common glass, and hence it is now generally used in lamp manufacture. Tungsten wire can be sealed into some of the commercial glasses which have a low coefficient of expansion.

197. Compensated pendulums. In a number of cases it is necessary to compensate for the change in dimensions produced by temperature changes. A well-known example of this is the case of clock pendulums. If the length of a pendulum is changed, its period is changed, and the clock runs slower or faster. Changes of temperature produce quite noticeable changes in the rate of a clock which has a simple metal pendulum. Clocks and watches which have in some way been made independent of temperature changes are said to be "compensated."

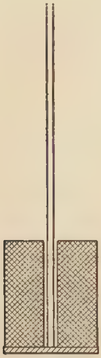


FIG. 141

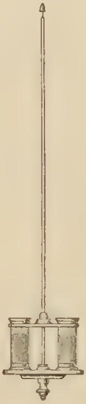


FIG. 142

One of the simplest devices is a pendulum with a well-seasoned wooden rod carrying a heavy zinc or lead cylinder at the bottom (Fig. 141). When the temperature rises, the rod expands slightly, lowering the cylinder. The expansion of the cylinder raises its center, for the cylinder is supported at the lower end. If the length of the cylinder is properly chosen, the two expansions will neutralize each other so far as changing the period of the pendulum is concerned. A better pendulum is one with a steel rod carrying at the lower end hollow cylinders of metal or glass partly filled with mercury (Fig. 142). When the temperature rises, the rod lengthens, thus lowering the mercury. But this is compensated by the expansion of the mercury. This is probably the best method of compensating a pendulum.

The principle of the "gridiron" pendulum can be understood by aid of Fig. 143. Rods *i*, *e*, and *d* are of iron, while rods *cc* are of brass. A rise in temperature causes all these rods to lengthen. The expansion of the iron tends to lower the heavy bob, but the expansion of the brass rods tends to raise it. Brass expands about 50 per cent more than iron, so that it is possible by a combination of iron and brass rods to compensate the pendulum for temperature changes.

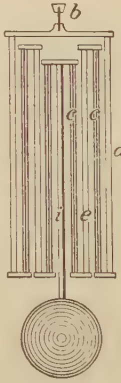


FIG. 143

198. Compensated balance wheels. The balance wheels of watches or of some types of clocks are torsion pendulums. The period of the balance wheel depends on the elastic force of the hairspring and also on the mass of the wheel and the distance of the mass from the axis (the moment of inertia of the wheel). Expansion of the wheel moves the mass farther from the axle and tends to increase the period. The rise of temperature weakens the elastic force of the hairspring; and this too tends to increase the period. Both these factors are compensated by making the rim of two different metals and by mounting it as shown in Fig. 144. When the temperature rises, the differential expansion (sect. 194) causes the free parts of the rim to move inward. Usually small metal screws are inserted symmetrically in the rim of the wheel. By placing the screws at different places and screwing them in or out, (1) the wheel can be balanced (that is, the center of gravity brought to the center of the axis), (2) the wheel can be given the proper period of vibration, and (3) the masses can be given the proper distribution so that the wheel is very accurately compensated for temperature changes.

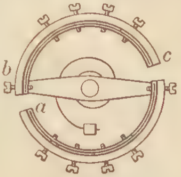


FIG. 144

In some modern watches and chronometers a simple brass balance wheel is used with a hairspring of a nickel-steel alloy. The elasticity of this alloy *increases* with an increase of temperature. The effect of the increase in elasticity of the spring tends to compensate the effect of the increase in the dimensions of the balance wheel. It is claimed that the compensation is better than that obtained with the more expensive balance wheel which is represented in Fig. 144.

199. Volume expansion. *Linear expansion* refers to the change in the length of any dimension of a body In all cases, except in crystalline substances, the rate of expansion of bodies with rise of temperature is the *same in all directions*. This is well shown by an experiment with a ring and a ball made of the same material. If the ring snugly fits the ball when they are both cold, it will fit snugly for all temperatures, provided they are both at the same temperature. The important fact to note is that any one dimension, any one measured distance, increases at the same rate as all other linear dimensions. For example, the diameter of the hole increases at the same rate as the diameter of the ball.

In the measurement of the changes in volume produced by rise of temperature, the following expression is useful:

$$\text{Change in volume} = \beta v_0 t, \tag{4}$$

where v_0 is the original volume, t the change of temperature, and β the coefficient of volume expansion.

If v_t is the volume at the higher temperature, the change of volume is

$$v_t - v_0 = \beta v_0 t,$$

or

$$v_t = v_0 + \beta v_0 t.$$

Hence

$$v_t = v_0(1 + \beta t). \tag{5}$$

The law for the change in volume of a substance is therefore similar to that for linear expansion. However, it is not accurate for very large temperature changes.

It can easily be proved that the value of the coefficient of volume expansion of any substance is very nearly three times the coefficient of linear expansion of that same material, or

$$\beta = 3\alpha.$$

By this relation the coefficients of volume expansion can be found from the coefficients of linear expansion given in section 195.

The following table gives the coefficient of volume expansion for a few liquids.

COEFFICIENTS OF VOLUME EXPANSION FOR DEGREES CENTIGRADE

Alcohol (ethyl)	.00110	Pentane	.00159
Mercury	.00018	Toluene	.00109
Paraffin oil	.00090	Water (15° – 100° C.)	.00037

200. Expansion of liquids; relative expansion. In the observation of the changes in volume that take place when a liquid is heated, one must usually take account of the expansion of the container. What is actually observed, for example, when a flask of water is heated, is the difference between the expansion of the flask and that of the water. This is often referred to as *relative expansion.* The effect of the expansion of the container can readily be shown in the case of a flask fitted with a stopper through which an open glass tube passes (Fig. 145). The flask is filled until the water stands in the tube above the stopper. When hot water is poured over the flask, or heat is quickly supplied in some other way, at first the liquid in the tube *falls*, indicating an *apparent contraction.* The explanation will readily occur to the student. The observed effect is due to the increase in the volume of the flask, which is heated first.

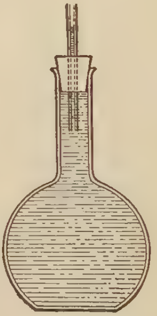


FIG. 145

In most cases of the expansion of liquids one is concerned only with the relative expansion of the liquid and its container. The working formula, or rule, for this is

$$v_t = v_0[1 + (\beta_1 - \beta_2)t], \qquad (6)$$

where v_t is the final *apparent* volume, v_0 the initial volume, β_1 the coefficient of volume expansion of the liquid, and β_2 the coefficient of volume expansion of the container.

201. Measurement of absolute expansion of a liquid. An ingenious method which does not depend on the expansion of the container was developed by Dulong and Petit and improved by Regnault. The principles involved can readily be understood by the aid of Fig. 146 (the actual details of the apparatus were much more complicated than this simple explanation indicates). The U-shaped tube *AB* is filled nearly full of the liquid to be tested. One arm is surrounded by ice water, and thus kept at a constant temperature of 0° C. The jacket around the other arm is filled with steam or some hot liquid at a temperature of, say, t. The effect of heating one column is to make its density less than that of the other. Since the liquid on one side is denser than that on the other, the liquid will not stand at the same height in the two arms. However, the pressure in the tube at the bottom must be the same on both sides. According to section 24 the pressure at the bottom of the cold column will be

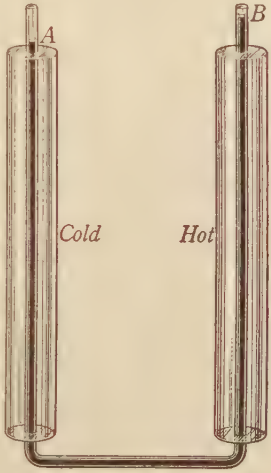


FIG. 146

$$p = h_0 d_0 g,$$

where h_0 is the height of the cold column, d_0 its density, and g the acceleration of gravity.

At the bottom of the warm column the pressure is given by

$$p = h_t d_t g,$$

where h_t is the height of the warm column and d_t its density.

Since the two pressures are equal,

$$h_0 d_0 g = h_t d_t g,$$

or

$$\frac{d_0}{d_t} = \frac{h_t}{h_0}.$$

But for equal masses the ratio of the densities is inversely that of the volumes, or

$$\frac{d_0}{d_t} = \frac{v_t}{v_0} = \frac{h_t}{h_0}.$$

From the law of expansion (equation (5)),

$$\frac{v_t}{v_0} = 1 + \beta t.$$

Hence

$$\frac{h_t}{h_0} = 1 + \beta t. \tag{7}$$

The heights h_t and h_0 and the temperature t can be measured. Hence the value of β can be computed.

202. Expansion of water; maximum density. The volume of pure water changes in the range from 0° C. to 10° C. in an unusual and noteworthy way. When water at 0° C. is warmed, it *contracts*, and continues to contract until it has reached a temperature of about 4° C. Above that temperature it expands. At a temperature a little over 8° C. it has the same density as at 0° C. Thus water has its *maximum* density at about 4° C. This unusual behavior is shown graphically in Fig. 147, where vertical heights show the relative volumes of a given sample of water at different temperatures t.

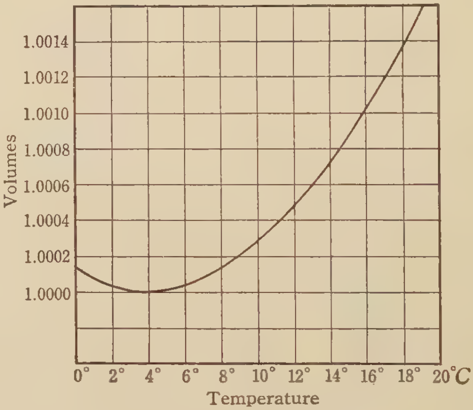


FIG. 147. The expansion of water

An interesting experiment which shows that the maximum density occurs at 4° C. is due to Hope. The glass vessel A (Fig. 148) is filled with water at room temperature, and the two thermometers T_1 and T_2 are read. A freezing mixture is then placed in the jacket CD. The thermometer T_2 begins to fall almost immediately, and keeps on falling until it reaches a temperature of about 4° C., when it stops. The temperature as indicated by thermometer T_1 remains practically constant during this time. But soon after T_2 reaches 4° C., T_1 begins falling, and continues to fall until it reaches 0° C. and the water begins to freeze. The thermometer T_2 stays at 4° C. during the time T_1 is falling. It is left as an exercise for the student to explain why the temperatures change in this way, and why this experiment shows that the temperature of the maximum density of water is 4° C.

This curious behavior of water is of great importance in nature. Were it not for this fact, ponds and lakes would much more easily freeze to the bottom, with a destruction of all animal life in them. In Lake Superior the temperature of the water at depths below 240 feet stays at about 4° C. all the year round.

The temperature of the maximum density of water is changed by salts in solution; for example, the maximum density of ocean water occurs a degree or two below 0° C., at about its freezing point.

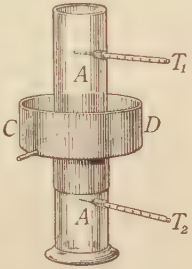


FIG. 148

The student can see why a thermometer made with water as the fluid would be useless below 10° C. (The rate of expansion does not become constant even up to 100° C.)

203. Forces involved in expansion. If a bar is heated but at the same time prevented from expanding, the force necessary to prevent expansion is the same as that required to compress the bar by an amount equal to the expansion produced by the change in temperature. For example, if the change of temperature is sufficient to cause a steel rail to elongate one tenth of an inch, the force required to prevent this expansion is the force necessary to shorten the rail one tenth of an inch by compression. On many street railway tracks the steel rails are welded together, and expansion is prevented by especially strong construction of the bed. It requires an enormous force to compress ice or water; hence when water freezes, it takes an enormous force to prevent the expansion. The destructive force of freezing water is well known.

204. Expansion of gases. The expansion of a gas is large compared with the expansion of liquids or solids. This can be shown by some such apparatus as that of Fig. 149. The flask contains air, and the U-shaped tube colored water. When the flask is heated,—for instance, by holding a burning match near it,—a large expansion of the air is clearly shown.

If one had air or any other gas in a cylinder with a movable piston (Fig. 150), the behavior of the gas could be studied under several different circumstances. Two of these are of interest here. A weight which is placed on the piston compresses the gas until it has a certain definite pressure. If the gas is now heated, it expands, raising the piston; but the pressure of the gas remains constant. This is a case of *expansion at constant pressure.* The other case is one in which the piston is kept from moving; hence the *volume* of the gas is kept *constant.* In this case the pressure of the gas changes with changes of the temperature. These two special cases will at first be discussed separately.

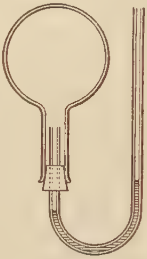


FIG. 149

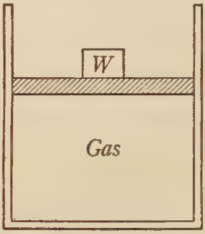


FIG. 150

205. Expansion at constant pressure. It is found by experiment that when the pressure is kept constant the volume v_t of any gas at a temperature t is related to v_0, the volume at 0° C., by the equation

$$v_t = v_0(1 + \tfrac{1}{273} t). \tag{8}$$

The fraction 1/273 is the coefficient of expansion of a gas at constant pressure.

206. Changes in pressure at constant volume. When the gas is kept at constant volume, the following law is true:

$$p_t = p_0(1 + \tfrac{1}{273} t), \tag{9}$$

where p_t is the pressure of the gas at a temperature t, and p_0 is the pressure at 0° C. The fraction 1/273 is in this case the coefficient of pressure change at constant volume.

Among the striking features of these laws is the fact that the coefficient of expansion at constant pressure is very nearly the same for all gases, and that this coefficient is equal (at least, nearly so) to the coefficient of pressure change at constant volume. It can be shown by relatively simple reasoning that for any gas which accurately obeys Boyle's law (sect. 45) these two coefficients must be equal. Except where unusual accuracy is required, these two coefficients can be regarded as numerically identical, an assumption which was made in equations (8) and (9) and which will be made in all that follows on this subject.

207. Summary of gas laws. We had, in section 45, Boyle's law, which states that

$$\text{For constant temperature, } pv = \text{a constant.} \qquad (10)$$

In the last two sections we had equation (8), stating that

$$\text{For constant pressure, } v_t = v_0(1 + \tfrac{1}{273}t),$$

and equation (9), stating that

$$\text{For constant volume, } p_t = p_0(1 + \tfrac{1}{273}t).$$

208. General gas law. The three laws stated in the last section can be combined into a single law of which the three are special cases. This *general* law is stated thus:

$$pv = p_0 v_0(1 + \tfrac{1}{273}t). \qquad (11)$$

Before going farther the student should be certain that he understands the meaning of the variables p, v, and t in equation (11). The gas can be in any possible condition as regards pressure, volume, and temperature. But whatever it is, these three letters refer to *simultaneous* values of the *pressure*, *volume*, and *temperature*. In this equation the pressure and volume may be measured in any units, but the temperature t must be measured in centigrade degrees.

The student can easily see by the following processes that the three laws given by equations (8), (9), and (10) are special cases.

Equation (10), Boyle's law, is true only when the *temperature* does not change. If, in equation (11), t is regarded as constant, all the factors on the right-hand side are constants, and the equation takes the form of equation (10).

If the *pressure* is kept constant (that is, if p is always kept equal to p_0, the pressure it would have at 0° C.), then the p's can be canceled from equation (11), and it reduces to equation (8). Or, if the *volume* is kept constant (that is, if v is always equal to v_0), the v's can be canceled from equation (11), and it becomes identical with equation (9).

Hence equations (8), (9), and (10) are special cases of the more general law given by equation (11).

The general law as stated in equation (11) becomes, by a slight algebraic change,

$$pv = \frac{p_0 v_0}{273}(273 + t).$$

The fraction on the right is a constant for a given mass of gas, and may be represented by R; that is,

$$R = \frac{p_0 v_0}{273}.$$

If we let

$$T = 273 + t,$$

equation (11) reduces to the simple form

$$pv = RT. \tag{12}$$

This is a common form of the *general gas law*. There is a simple interpretation of T. It represents temperature measured on a scale which has as the zero point $-273°$ C. This is commonly called the gas-temperature scale. This scale is practically the same as the absolute, or Kelvin, scale. The reason for the use of the gas scale is obvious, for the general law as stated in equation (12) is much simpler than in equation (11).

The general gas law is often stated by the equation

$$pv = mR'T, \tag{13}$$

where m is the *mass* of the gas, and the product mR' is equal to the constant R of equation (12).

209. Examples. 1. The pressure of a gas is 400 cm. of mercury at 100° C. What is the pressure at 50° C. if the volume is kept constant and equal to 500 cc.?

We may use p_1, v_1, and T_1 to refer to the pressure, volume, and temperature in the first state of the gas, and p_2, v_2, and T_2 to refer to the same quantities for the second state. Since the gas law holds for all states,

$$p_1 v_1 = RT_1 \tag{14}$$

and

$$p_2 v_2 = RT_2. \tag{15}$$

If equation (15) is divided by equation (14), we obtain

$$\frac{p_2 v_2}{p_1 v_1} = \frac{T_2}{T_1}. \tag{16}$$

Since in our problem $v_2 = v_1$,

$$\frac{p_2}{p_1} = \frac{T_2}{T_1}, \tag{17}$$

which is true for any gas kept at *constant volume.*

$$T_2 = 273 + 50 = 323° \, K,$$

$$T_1 = 273 + 100 = 373° \, K,$$

and $p_1 = 400$ cm. of mercury.

Substituting these values in equation (17),

$$\frac{p_2}{400} = \frac{323}{373},$$

$$p_2 = 346.4 \text{ cm. of mercury.}$$

2. The volume of a gas at atmospheric pressure is 200 cc. when the temperature is 20° C. What is the volume when the temperature is 50°, the pressure being kept constant?

Using equation (16) and remembering that for this case $p_2 = p_1$,

$$\frac{v_2}{v_1} = \frac{T_2}{T_1}, \tag{18}$$

an expression always true when the *pressure is kept constant.* In the problem

$$v_1 = 200 \text{ cc.},$$

$$T_1 = 273 + 20 = 293° \, K$$

and $T_2 = 273 + 50 = 323° \, K.$

Hence
$$\frac{v_2}{200} = \frac{323}{293}$$

and $v_2 = 220.5$ cc.

In any problem in which all three quantities change, equation (16) will be found very useful. It should be obvious that any problem can be solved by this equation, provided that only one of the six quantities involved is unknown.

Equations (17) and (18) may be stated as follows:

1. If the volume of a gas does not change, the pressure is proportional to the temperature measured on the Kelvin scale.

2. If the pressure of a gas does not change, the volume is proportional to the temperature measured on the Kelvin scale.

As far as their usefulness is concerned, the gas laws in the form of equations (16), (17), and (18) are better for practical work than equations (8), (9), and (11).

210. Limit of application of gas laws. The student must not assume that these gas laws apply down to very low temperatures. At sufficiently low temperatures and moderate pressures gases become liquids, and certainly do not obey these laws. This remark is made because it is often stated that since the pressure of a gas is proportional to the temperature T, its pressure will be zero when $T = 0$; that is, at $-273°$ C. Sometimes the statement is made that the volume will also be zero at this temperature, an obviously absurd statement. The student must remember that these laws are for gases, and are true only within certain limits. When a gas is cooled to near its liquefying temperature, it departs widely from these laws.

211. The gas thermometer. For measuring temperatures by a gas thermometer, either one of two principles is used:

1. *The constant-pressure thermometer.* The volume of the gas in the thermometer is measured at some known temperature, and again at the temperature which is to be determined, the pressure being kept constant. By equation (18) we have

$$\frac{v_2}{v_1} = \frac{T_2}{T_1}.$$

Since both volumes have been measured and one temperature is known, the other temperature can be computed by this equation.

2. *The constant-volume thermometer.* The pressure of the gas in the thermometer is measured at some known temperature, and again at the unknown temperature, the volume being kept constant. In this case we have, by equation (17),

$$\frac{p_2}{p_1} = \frac{T_2}{T_1}.$$

Since both pressures are known, and one of the temperatures, the other temperature can be computed.

The *standard* scale of temperature generally used for scientific purposes is that given by the constant-volume thermometer. The essential parts of a simple form of constant-volume thermometer are shown in Fig. 151. The bulb A, made of glass, quartz, porcelain, or some similar material, contains the gas under test. When the temperature of the gas is changed, the level of the mercury in column B is brought back to the same point as before by raising or lowering a reservoir of mercury which is connected to B and C by a flexible tube. The pressure of the gas in A is readily measured, for it is equal to barometric pressure plus the pressure that is due to the difference in the heights of the mercury columns C and B. Temperatures are obtained, as already explained, by a simple computation involving equation (17).

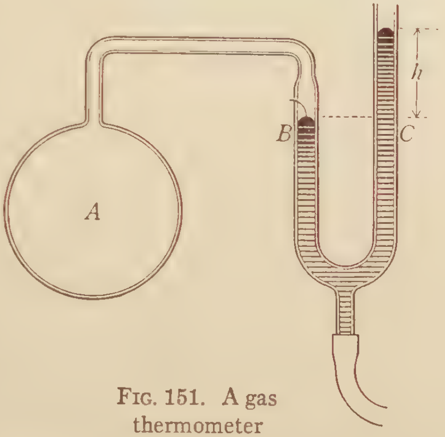


FIG. 151. A gas thermometer

212. Summary. The law of *linear expansion* is given by equation (3),

$$l_t = l_0(1 + \alpha t),$$

where α is the *coefficient* of linear expansion. A table of numerical values of α is given in section 195. The student should be able to define this coefficient by remembering that the total expansion is given by the product $l_0 \alpha t$.

The law of *volume expansion* is, by equation (5),

$$v_t = v_0(1 + \beta t),$$

where the coefficient of volume expansion β is approximately three times the coefficient of linear expansion. The value of β for a few liquids is given in section 199.

In the case of a liquid in a container which also expands, the *apparent* volume of the liquid at a temperature t is, by equation (6),

$$v_t = v_0[1 + (\beta_1 - \beta_2)t],$$

where β_1 is the coefficient for the liquid and β_2 that of the container.

The maximum density of water occurs at 4° C.

The general gas law is, by equation (12),

$$pv = RT,$$

where T is temperature measured on the gas scale. The gas scale is practically identical with the Kelvin, or absolute, scale. The temperature T is the temperature in centigrade plus 273. The following are convenient working equations:

(17), for constant volume, $\frac{p_2}{p_1} = \frac{T_2}{T_1}$.

(18), for constant pressure, $\frac{v_2}{v_1} = \frac{T_2}{T_1}$.

(16), for all quantities variable, $\frac{p_2 v_2}{p_1 v_1} = \frac{T_2}{T_1}$.

PROBLEMS

1. How much will the length of 200 ft. of copper wire change for a change in temperature of 50° C.? (The coefficient is .000016.)

2. An iron steam-pipe line is 100 ft. long at 20° C. What will be the increase in its length when steam at 105° C. passes through it? (The coefficient of expansion is .000012.)

3. Calculate the change in length of a steel bridge 300 ft. long between a winter minimum of $-30°$ C. and a summer maximum of $+40°$ C. if the coefficient for steel is .000012.

4. A steel scale which was correct at 0° C. was used to measure a certain distance. The reading, 84.12 cm., was taken at a temperature of 30° C. What was the error? (The coefficient of steel is .000012.)

5. How long must a brass bar be in order that its change of length may be the same as that of a steel bar 100 cm. long? (The coefficient of steel is .000012; of brass, .000018.)

6. A steel ball displaces 400 cc. of water at 0° C. How many cubic centimeters will be displaced at 30° C.?

7. A steel rod has its temperature raised 20° C. How much compressional stress must be applied to this rod in order to prevent it from expanding? (The coefficient of expansion is .000012, and Young's modulus is 28×10^6 lb./sq. in.)

8. A certain mass of gas has a volume of 1000 cc. at 50° C. What will be the volume if the temperature is raised to 90° C., the pressure remaining constant?

9. Air passing through a hot-air furnace has its temperature changed from 15° C. to 60° C. What is the percentage increase of its volume?

10. A steel vessel holds 500 gm. of mercury at 0° C. How many grams will it hold at 30° C.? (The linear coefficient of steel is .000012; the volume coefficient of mercury is .00018, and its density is 13.60 at 0° C.)

11. A glass vessel contains 200 cc. of alcohol, and is full at a temperature of 15° C. How many cubic centimeters will spill over when the temperature is raised to 40° C.? (The volume coefficient of glass is .00002; of alcohol, .00110.)

12. The air in an automobile tire had a pressure of 80 lb./sq. in. (above atmospheric pressure, which is 15 lb./sq. in.) at a temperature of 20° C. What will be the pressure in the tire when, from fast driving, its temperature is raised to 60° C.? (Assume the volume to be constant.)

13. A constant-volume gas thermometer was used to measure temperature. At 0° C. the observed pressure was 86.2 cm. At another temperature the pressure was 98.8 cm. Compute the temperature.

14. A gram of air occupies 774 cc. at 0° C. and at a pressure of 76 cm. of mercury. What will be its volume at 20° and at a pressure of 73 cm. of mercury?

15. A certain mass of dry air at a temperature of 20° C. and a pressure of 74 cm. of mercury has a volume of 500 cc. Find the pressure at a temperature of 30° C. when the volume is 600 cc.

CHAPTER XVIII

QUANTITY OF HEAT

Quantity of heat, 213. Heat a form of energy, 214. Units for measuring quantity of heat, 215. Method of mixtures, 216. Calorimetry, 217. Thermal capacity, 218. Specific heat, 219. Experimental determination of specific heat, 220. Summary, 221.

213. Quantity of heat. A comparatively simple experiment will bring out an important fact. To perform this two Bunsen burners giving the same kind and type of flame are used. One of them is placed under a suspended block of iron weighing about a pound, and the other is at the same time placed under a pound of water in a thin-walled vessel. In a few minutes the iron will be heated above the boiling point of water, as can be shown by dropping a little water on it, while the equal mass of water will still be cool enough for one to hold his fingers in it. If we assume that the two, the iron and the water, receive *heat* at the *same rate*, it is obvious that the *increase in temperature* of the two is not the same thing as the *quantity of heat* they have received. Other experiments might be cited to support this. For example, if a pan of crushed ice is placed on a hot stove, the mixture of ice and water will remain at a constant temperature of 0° C. until most of the ice is melted. In this case, adding *heat* did not produce at once an increase in *temperature*. Hence it is necessary to distinguish between *quantity* of heat and *temperature*. The distinction should be seen clearly before the student has completed this chapter.

214. Heat a form of energy. The old theory was that heat was a material substance, called *caloric*, which in some way was able to penetrate bodies. A hot body contained a large amount of it, while a cold body contained a smaller amount. It was necessary to assume that this substance was not like ordinary matter; for example, it possessed neither inertia nor weight.

The modern theory regards heat as a *form of energy*. According to this theory the molecules, atoms, and electrons in all bodies are in a continual state of vibration. In a hot body these small particles have relatively large amounts of kinetic energy, while in a cold one they have less. From this point of view it is easy to understand how friction develops heat in a body, for when one body slides over another the molecules on the two surfaces are given an increased vibration. It is also important to look at this from the standpoint of the principle of conservation of energy. Because the small particles are given motion, and hence kinetic energy, work must be done to slide one body over the other. Mechanical energy *apparently* disappears. But the "lost" energy is now kinetic energy of the small subdivisions of matter,—a form of energy we call heat.

It will appear later that what we now call *quantity* of heat is really the *energy* of these small subdivisions of matter.

215. Units for measuring quantity of heat. Three different units are commonly used in measuring *quantity of heat*. They are defined as follows:

The calorie, or gram-calorie, is equal to the amount of heat required to change the temperature of 1 gram of water 1 degree centigrade.

The kilogram-calorie is equal to the amount of heat required to change the temperature of 1 kilogram of water 1 degree centigrade.

The British thermal unit (B.T.U.) is equal to the amount of heat required to change the temperature of 1 pound of water 1 degree Fahrenheit.

To aid in fixing these units in mind the student should mentally verify the following calculations:

The quantity of heat required to raise 1 gram of water from 10° C. to 30° C. is 20 calories.

The quantity of heat required to raise 400 grams of water from 10° C. to 40° C. is 12,000 calories, or 12 kilogram-calories.

The quantity of heat required to raise 5 pounds of water from 40° F. to 65° F. is 125 British thermal units.

The preceding definitions are only approximations, although they are sufficiently close for all except very accurate work. More precise definitions should state the exact temperature range. For example, the precise definition of the calorie may be stated as the quantity of heat necessary to change 1 gram of water from 15° C. to 16° C. The reason for this explicit statement is that it is not true that the same amount of heat is required to change 1 gram of water 1 degree whatever the initial temperature may be. The amount is slightly different for different ranges.

216. Method of mixtures. The solution of the following problem brings out a principle which has a very wide use in quantitative work in heat.

Problem. 200 gm. of water at 10° C. are mixed with 500 gm. of water at 30° C. Compute the temperature of the mixture.

Solution. In this mixture 200 gm. of water gain heat, while 500 gm. lose heat. If the mixing takes place under such circumstances that heat is not gained from, or given up to, other bodies in the neighborhood, the heat gained by the one must equal that lost by the other part (if it is granted that heat is a form of energy, this follows from the principle of conservation of energy).

The heat gained is $200(t-10)$ cal., and the heat lost is $500(30-t)$ cal., where t is the temperature of the mixture. Equating these,

$$200(t-10)=500(30-t),$$

or

$$t=24.3^\circ \text{ C.}$$

The important principle is expressed by saying that the number of calories gained by one part is equal to the number of calories lost by the other part.

217. Calorimetry. Since the usefulness of coal is proportional to the quantity of heat that it gives off when burned, it is often bought and sold on the basis of the number of British thermal units that 1 pound gives off. The heat from a sample can be measured by an apparatus in which all the heat given up by the coal goes into a water jacket which surrounds the combustion chamber. If the weight of the water and the change of temperature are known, the number of British thermal units can be computed. This is one example of many cases in which it is important to measure the quantity of heat. That branch of the subject that deals with heat measurements is called *calorimetry.*

218. Thermal capacity. The results of the experiment referred to in section 213, of heating equal masses of iron and water, are often stated by saying that water and iron do not have equal *thermal capacities.*

There is another simple experiment which shows a difference in *thermal capacities* of different bodies: Small balls of copper, iron, tin, aluminum, etc., each about 2 centimeters in diameter, are heated in oil to a temperature somewhat above that of boiling water and then dropped on a thin cake of paraffin. The copper and iron balls may melt their way through the wax and fall to the table below, while the others sink to different depths in the wax. The different balls, as judged by the amount of wax they melt, give up different quantities of heat in cooling from the same high temperature down to the temperature of the melting wax. They are said to have different *thermal capacities.*

The thermal capacity of any body (for example, that of one of the balls just referred to) *is equal to the number of calories required to change its temperature 1 degree centigrade.* This applies equally to the case of raising and that of lowering the temperature of the body, for the number of calories required to raise the temperature 1 degree is the same as the number of calories given up when the temperature falls 1 degree.

219. Specific heat. *The specific heat of any substance is equal to the number of calories required to raise the temperature of 1 gram of the substance 1 degree centigrade.* For example, an aluminum vessel has a mass of 500 grams. The specific heat of aluminum is about 0.21 calories per gram per degree. Since the specific heat is the number of calories necessary to raise the temperature of 1 gram of aluminum 1 degree, the number of calories required to raise the entire vessel 1 degree is $500 \times 0.21 = 105$ calories per degree, which is the thermal capacity.

The thermal capacity of a body is equal to the specific heat of the material multiplied by the mass of the body.

The following expressions will be found useful. The student should think about them until he knows why they are true.

The number of calories required to raise the temperature of a body from t_1 to t_2 degrees centigrade

$$= \text{thermal capacity of body} \times (t_2 - t_1)$$
$$= \text{mass of body} \times \text{specific heat} \times (t_2 - t_1).$$

From the definitions of a calorie and of specific heat it should be clear that the specific heat of water is equal to 1.

220. Experimental determination of specific heat. A copper vessel, mass 100 grams, contains 200 grams of water at 15.0° C. A mass of copper, of 300 grams, is heated to a temperature of 100° C. in a steam bath and then suddenly plunged into the vessel of cold water. The final temperature of the mixture is observed to be 24.8° C. From these data the specific heat of the copper can be computed.

Let s be the specific heat of the copper. The copper vessel and the water contained in it are increased in temperature from 15.0° C. to 24.8° C. The number of calories they have gained is $100\,s\,(24.8 - 15.0) + 200\,(24.8 - 15.0)$. The 300 grams of copper cool from 100° C. to 24.8° C. and lose $300\,s\,(100 - 24.8)$ calories. The heat received from or given to the room by the vessel and contents during the experiment may be neglected, for the temperature of the vessel was, on the average, probably not far from room temperature. If this is done, we may equate the number of calories gained to those lost, and thus get

$$100\,s\,(24.8 - 15.0) + 200(24.8 - 15.0) = 300\,s(100 - 24.8).$$

From this $$s = \tfrac{1960}{21580}$$
$$= .091 \text{ cal. per gram per degree.}$$

Hence it requires less than one tenth of a calorie to change the temperature of 1 gram of copper 1 degree centigrade. It follows from the definition of a calorie that this is less than the heat required to change the temperature of one tenth of a gram of water 1 degree.

The specific heat of all common solids and liquids is much less than unity, the specific heat of water. The fact that the specific heat of water is relatively so great is the reason why large masses of water have such a moderating effect on climate. Islands surrounded by extensive masses of water do not have the same extremes of temperature as are observed in the interior of a continent. The low winter temperatures and high summer temperatures observed in the central states of the United States are due in large measure to the absence of large bodies of water. Leningrad and London differ in latitude by only a few degrees, yet there is a great difference in their climates, Leningrad being hotter in the summer and much colder in the winter.

TABLE OF SPECIFIC HEATS*

(In Calories per Gram per Degree Centigrade)

Alcohol (ethyl)	0.55	Lead	0.031
Aluminum	0.21	Mercury	0.033
Copper	0.093	Nickel	0.11
Glass (crown)	0.16	Oil (paraffin)	0.51
Glass (flint)	0.12	Paraffin wax	0.7
Glycerin	0.58	Silver	0.056
Gold	0.030	Steam	0.46
Ice	0.50	Tin	0.05
Iron	0.11	Zinc	0.092

221. Summary. It is important to distinguish between *temperature* and *quantity of heat.* The modern theory is that heat is a form of energy, and that *quantity* is the amount of energy that exists in the form of heat.

Three units for the measurement of quantity of heat are used: the *calorie*, equal to the amount of heat required to change the temperature of 1 gram of water 1 degree centigrade; the *kilogram-calorie*, equal to the amount of heat required to change the temperature of 1 kilogram of water 1 degree centigrade; the *British thermal unit* (B.T.U.), equal to the amount of heat required to change the temperature of 1 pound of water 1 degree Fahrenheit.

The *thermal capacity* of any body is equal to the number of calories required to change its temperature 1 degree centigrade.

The *specific heat* of a substance is equal to the number of calories required to raise the temperature of 1 gram of the substance 1 degree centigrade.

The specific heat of water is larger than that of other common substances. This fact plays an important part in natural phenomena.

---

*For specific heats of gases see section 230.

PROBLEMS

1. What will be the temperature of 400 gm. of water, initially at 15° C., when 2 kg.-cal. of heat are added?

2. What is the thermal capacity of 250 gm. of water in calories per degree?

3. 500 gm. of water at 15° C. are mixed with 200 gm. at 60° C. Compute the resulting temperature.

4. What is the specific heat of copper if the thermal capacity of a mass of 1.2 kg. is 112 cal. per degree?

5. 500 gm. of aluminum at 90° C. are mixed with 200 gm. of water at 10° C. What is the resulting temperature if the specific heat of aluminum is 0.21 cal. per gram per degree?

6. A silver dish of 50 gm. contains 200 gm. of water at 16° C. A piece of silver of 65 gm. is heated to 100° C. and then plunged into the water. What is the resulting temperature? (The specific heat of silver is 0.056 cal. per gram per degree.)

7. A mass of 150 gm. of zinc at 100° C. is dropped into a 60-gram glass beaker (specific heat 0.16 cal. per gram per degree) containing 70 gm. of water at 12° C. The resulting temperature is 25° C. Compute the specific heat of zinc.

8. Into 12 kg. of water at 30° C. are dropped 1 kg. of iron at 100° C. and 1.2 kg. of zinc at 60° C. Find the resulting temperature. (The specific heat of iron is 0.11 cal. and that of zinc 0.092 cal. per gram per degree.)

9. 150 gm. of copper (specific heat 0.093 cal. per gram per degree) was dropped into 400 gm. of water at 10° C. The copper had been heated to a temperature of 99° C. in a water bath. When moved over, it carried with it 2 gm. of the hot water. Compute the resulting temperature.

CHAPTER XIX

WORK AND HEAT

Reasons for the overthrow of the caloric theory, 222. Joule's experiment, 223. The mechanical equivalent of heat, 224. Examples, 225. Compression and expansion of gases, 226. Adiabatic compression and expansion, 227. Other illustrations of the heating and cooling of gases by compression or expansion, 228. Two important cases; free expansion, 229. Two specific heats of gases, 230. Heats of combustion, 231. Thermodynamics, 232. The first law of thermodynamics, 233. The second law of thermodynamics, 234. Dissipation and degradation of energy, 235. Summary, 236.

222. Reasons for the overthrow of the caloric theory. The reasons for the abandonment of the old *caloric* theory and the acceptance of the modern theory may be classified into three groups:

1. The concept of a substance like the caloric leads to inconsistencies. If heat is a substance, one would expect that in an isolated group of bodies the total amount of this substance would remain constant. In many cases it does, but in some it does not. For example, in many well-known cases heat is generated by friction. The caloric theory attempted to explain this by saying that the caloric was already in the body, and that the friction rubbed it out. According to this idea there should be less caloric left in a body which had been rubbed. Count Rumford (1753–1814) was one of the first to challenge this. While in charge of the manufacture of cannon at Munich he was greatly impressed by the immense quantity of heat developed in the boring process. Using a blunt drill, he generated large amounts of heat. He even boiled water which had been placed in a surrounding jacket. He argued that if heat were a substance, no body could have the apparently unlimited supply that could be brought out through friction. Another fact worth mentioning is that the shavings or borings have the same specific heat as the solid materials; that is, both the shavings and an equal mass of the solid metal give out the

same amount of heat when cooled through the same range of temperature. None of the thermal capacity is rubbed out by the friction.

Sir Humphry Davy found that rubbing two pieces of ice together in a vacuum, where heat from other bodies could not reach them, melted the ice,—a process requiring heat. Shaking or churning water was found to warm it. Heat is generated in iron or lead when it is hammered. If one takes a small piece of iron wire, and bends it back and forth vigorously as if trying to break it, the wire will get appreciably warm at the bent part. There is internal friction in the wire, and work is done in bending it back and forth. Other instances, such as a hot-box on a railroad car, will occur to the student. To generalize, it was found that in all those cases where work was done on a body without any visible changes in the potential or kinetic energy of the body, heat was generated. In other words, *when mechanical energy disappeared, heat always appeared.* This the caloric theory could not satisfactorily explain.

2. The second group is based on a remarkable fact discovered in connection with the generation of heat by mechanical work. This fact was that the number of calories of heat generated was *always proportional* to the mechanical energy lost; or, what is the same thing, for every calorie of heat generated a certain definite number of ergs of mechanical energy disappeared. This was found true in all cases, in many experiments of different types.

An analogy at this point will be helpful. Suppose that both a yard and a meter stick are used to measure a large number of different things. The ratio of any length measured in yards to the length of the same thing measured in meters will always be equal to 1.09 approximately (this is the number of yards in 1 meter). This constant ratio of 1.09 is obtained because a yard and a meter are units for measuring the *same thing*; namely, *length.* If they measured different things, we should not expect to get this constant ratio. The observed constant ratio between ergs and calories indicates that these two units measure the same thing. If heat is a form of energy, we should expect a definite numerical ratio between ergs and calories, for in that case both are units for measuring energy.

3. The tendency in scientific work is toward unification. One of our great principles is the law of the conservation of energy. Experimentally it was found true in countless mechanical cases.

But when attempts were made to apply it to all cases, a serious stumblingblock was always obvious; for whenever there was any friction, energy apparently disappeared. If it is granted that heat is a form of energy, then energy is not destroyed through friction but is merely changed into another form of energy—energy of agitation of the electrons, atoms, and molecules. Assuming this, we no longer find an exception to the principle of the conservation of energy. It must, then, be true not only in mechanics but in all branches of science—true for everything, everywhere. The remarkable success with which this great principle has been used in the discovery of new facts and in the correlation of old ones gives strong support for the belief that it really is a *general* principle; if it is, heat must be a form of energy.

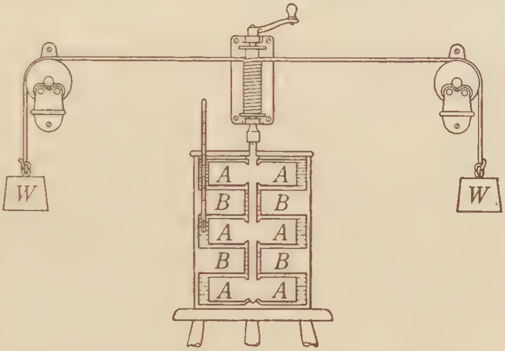


FIG. 152

223. Joule's experiment. We shall now take up one of the experimental methods which give the numerical relationship between *ergs* and *calories*. To one interested in the development of the theory of heat this experiment is of great significance, for it shows beyond all chance for controversy that the mechanical energy which disappears is proportional to the number of calories of heat generated, and hence that a definite number of ergs of work is required for each calorie of heat generated. To one who has accepted the fact that heat is a form of energy and that both ergs and calories are energy units this experiment is of interest because it gives a method for obtaining the number of ergs in 1 calorie.

The essential parts of Joule's apparatus were a copper vessel with a water-tight lid, a churning device consisting of a number of vanes (or paddles), which could be rotated in the water, and a number of stationary vanes placed between the moving ones. The moving vanes are shown at A, and the stationary ones at B, in Fig. 152. By means of these the water was churned, and the slight increase in the temperature of the vessel and its contents was carefully measured. If the mass of the water, vessel, vanes, etc. and the change in temperature are known, the number of calories given to the vessel and its contents can be computed. The work done in churning the water must also be measured. Joule used two falling weights to drive his churn. These were connected to the spindle of the churn by strings which ran over nearly friction-free pulleys. In falling, the weights turned the spindle and the moving vanes. After the weights had reached the floor they were lifted up again, and the process was repeated. In one experiment* the mass of the two weights was 26,300 grams, and each fell 160.5 centimeters twenty times in succession. The total number of ergs of work done was

$$20 \times 160.5 \times 26{,}300 \times 981 = 8.29 \times 10^{10}.$$

The rise in temperature of the vessel and its contents during the experiment was 0.3129° C. From the mass of the water, vessel, vanes, etc. and from the rise of temperature, computation showed that the number of calories added was 1977. If all the mechanical energy went into heat, then

$$1977 \text{ cal.} = 8.29 \times 10^{10} \text{ ergs.}$$

Hence

$$1 \text{ cal.} = 4.19 \times 10^{7} \text{ ergs.}$$

There were numerous corrections for which Joule made allowances in his final results; for example, the energy lost in friction in the pulleys, the kinetic energy the weights had when they were stopped by the floor, and the loss of heat from the vessel to the room. It is not necessary to give here the details of these corrections.

Similar experiments have been performed by many others, some of these being much more accurate than the original ones of Joule. The results are all in excellent agreement.

224. The mechanical equivalent of heat. If W is equal to the work (measured in *ergs*) used in generating heat, and H the heat developed (measured in *calories*), the experiments of Joule and others showed that the ratio W/H is a constant quantity. This constant quantity is usually designated by the letter J (the initial letter of *Joule*).

---

* Edser, Heat, p. 277.

Since
$$\frac{W}{H} = J,$$
$$W = JH. \qquad (19)$$

Or, The work done in ergs $= J \times$ calories of heat developed.

In this case J is equal to the number of ergs of work required to produce 1 calorie of heat. If the mechanical work, W, is measured in foot-pounds, and the heat, H, in British thermal units, J is equal to the number of foot-pounds of energy required to produce 1 British thermal unit.

The ratio J is usually called the *mechanical equivalent of heat.* Hence *the mechanical equivalent of heat is the number of units of mechanical energy required to produce 1 unit of heat.*

The following are accepted values of the mechanical equivalent:

$$\begin{aligned} 1 \text{ cal.} &= 4.18 \times 10^7 \text{ ergs} \\ 1 \text{ cal.} &= 4.18 \text{ joules} \\ 1 \text{ cal.} &= 3.09 \text{ ft.-lb.} \\ 1 \text{ B.T.U.} &= 778 \text{ ft.-lb.} \end{aligned}$$

The following values are often convenient in computations:

$$\begin{aligned} 1 \text{ joule} &= 0.239 \text{ cal.} \\ 1 \text{ ft.-lb.} &= 0.324 \text{ cal.} \\ 1 \text{ ft.-lb.} &= 1.286 \times 10^{-3} \text{ B.T.U.} \\ 1 \text{ B.T.U.} &= 0.252 \text{ kg.-cal.} \end{aligned}$$

225. Examples. Several types of problems are given here, to illustrate some of the computations that are possible.

1. A 200-pound weight falls 40 ft. How much heat is developed?

A 200-pound weight in falling 40 ft. loses potential energy of 200 × 40, or 8000 ft.-lb. If this is all converted into heat, the resultant number of calories is $8000/3.09 = 2590$ cal., or the number of British thermal units is $8000/778 = 10.3$ B.T.U.

2. A moderately active man should receive daily an average of about 3000 kg.-cal. of energy. He gets this energy from his food. To how many foot-pounds does this correspond?

$$3000 \times 3090 \text{ ft.-lb.} = 9.27 \times 10^6 \text{ ft.-lb.}$$

3. A lead bullet, the mass of which is m, hits a target with a velocity of 300 m. per second. Assuming that half of the heat developed goes into the target and the rest into the bullet, how much will the temperature of the bullet be raised?

The kinetic energy of the bullet is $\frac{1}{2}\,m(30{,}000)^2$ ergs $= 4.5\,m \times 10^8$ ergs. This is equal to $\frac{4.5\,m \times 10^8}{4.2 \times 10^7} = 10.7\,m$ calories. Half of this is $5.35\,m$ calories. Let t be the amount the temperature is raised. Since the specific heat of lead is 0.031, $0.031 \times m \times t = 5.35 \times m$, or $t = 5.35/0.031 = 173°$ C. It will be noticed that the mass m cancels out. This means that the rise in the temperature of the bullet, under the assumption made, is independent of its mass.

226. Compression and expansion of gases. It is a well-known fact that when a gas is compressed its temperature rises. A bicycle-tire pump gets hot at the place where the greatest compression of air occurs. Large air-compressors require special cooling devices. In one of the operations of a refrigerating machine, ammonia vapor is compressed, a process which produces considerable heat. Each of these cases may be regarded as an example of the principles which have just been stated, for they are cases of the conversion of mechanical energy into heat energy. When a gas is compressed, work is done on it, and mechanical energy apparently disappears.

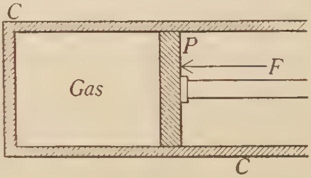


FIG. 153

It is worth while to consider more carefully this transfer of mechanical work into heat, and to try to form in our imaginations a fairly complete picture of the actual processes that take place in this conversion. To do this we shall consider a relatively simple case. In Fig. 153 is shown in section a cylinder CC with a movable piston P. Suppose that the closed part of the cylinder contains a gas; for example, air. When the force F moves the piston in, the gas is compressed and work is done. This work is done at an expenditure of energy from some external source. But energy is never destroyed. Where has this energy gone? By simple reasoning it will be shown that the kinetic energy of the molecules of the gas is increased by the motion of the piston.

The kinetic theory of gases (sect. 38) tells us that the pressure of a gas is due to the bombardment of the walls of the vessel by molecules, which move with remarkably large velocities. When the walls of the vessel are stationary, these molecules rebound with an unchanged speed; no speed nor energy is lost. But when the piston is moving in, these molecules, striking it, will rebound with greater velocity than they would if it were stationary. In order to understand this one has only to recall that a bat moving to meet a baseball will give the ball greater velocity and greater energy than if the bat were held stationary. Hence moving the piston in will increase the velocities and will therefore increase the average kinetic energy of the gas molecules. But an increase in the kinetic energy of the molecules of any substance means a rise in temperature. Hence we should expect the compression of a gas to heat it.

By the aid of quite similar reasoning the student can show that when a gas expands, the average kinetic energy of the gas molecules becomes less, and the gas becomes cooler. This is true, however, only in the case where a gas does work in expanding (see section 229).

227. Adiabatic compression and expansion. If compression or expansion of a small quantity of gas takes place slowly, the heating or cooling effect may not be noticeable, because the gas may give up heat to the vessel or bodies with which it is in contact or receive heat from them. The special case where the compression or expansion of a gas takes place under such circumstances that the gas does not receive heat from nor lose it to other bodies is called *adiabatic compression* or *adiabatic expansion.*

228. Other illustrations of the heating and cooling of gases by compression or expansion. One of the best experiments for the classroom is the so-called fire syringe. This is a narrow cylinder of glass which is closed at the lower end and which contains a well-fitting piston. If a small piece of cotton containing a very little carbon bisulfide is placed in the cylinder, an explosive gas will be formed. When the piston is driven down very quickly by a sudden blow of the hand, a flash of light will usually be seen in the glass tube. When the compression takes place very quickly, there is not time enough for the gas to give up an appreciable amount of heat to the walls of the tube; hence the gas is raised to a high temperature.

A simple apparatus can be used to show the effects of both compression and expansion. A large vessel (for example, a glass carboy) is fitted with a stopper through which two glass tubes run (Fig. 154). One of these tubes is connected to the pressure gauge *B*, which should contain some oil (an oil gauge being more sensitive than one containing a denser liquid). The other tube, *A*, is fitted with a rubber tube which can be attached to a compression pump or pinched off, as needed. Air is pumped into the carboy until a pressure of 20 or 30 centimeters of oil is obtained. This slight compression will warm the air; and after standing a little while it will cool, as shown by a decrease in its pressure. After the air in the carboy has cooled to room temperature, or approximately so, the tube *A* is opened, permitting the air in the carboy to expand until it reaches atmospheric pressure. This expansion cools the air in the carboy below room temperature. If the tube *A* is now pinched off, the gauge will show that the pressure in the carboy slowly rises; for the air in the carboy, being below room temperature, will receive heat from the room, and the rise of temperature will cause an increase in the pressure. This experiment thus shows that when the air is first compressed into the carboy, it is warmer than the room. Then, later, when the air in the carboy expands, it becomes colder than the room.

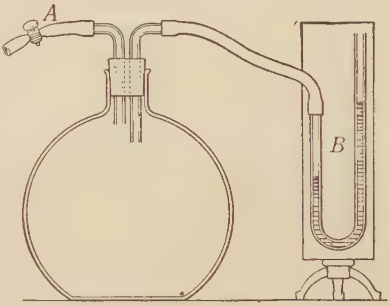


Fig. 154

Very interesting effects are shown in the movements of the atmosphere. Air descending from high altitudes, where the atmospheric pressure is small, will be compressed, and hence heated. The so-called chinook winds, which are quite common in some parts of America, are dry winds which are warmed by compression as they descend from high altitudes. The prevailing direction of the winds over the Rocky Mountains is from the west. As they pass over the mountains they descend on the eastern slopes and are warmed by compression. At certain seasons these winds become almost permanent, exerting an appreciable influence on the climate. This is especially noticeable in the Saskatchewan valley in Canada and in places in Montana and Wyoming. The air being dry, snow rapidly evaporates, and often the valleys are kept free from snow the greater part of the winter. The foehn wind of the Alps is similar in cause and character to our chinook wind. On the other hand, ascending currents of air are cooled by expansion. This fact plays such an important part in cloud formation and in rainstorms that it will be explained more fully later.

Almost everyone has marveled at the immense amount of heat that the sun radiates. Especially is this true when it is fully realized that the sun does this continuously, day after day, year after year, and that it has been doing this for millions of years. Where does all this energy come from? Modern theory indicates that the sun was at one time a mass of gas which filled an immense space. Under its own gravitational force this gas has slowly contracted, getting denser and denser, until its volume has reached the present size. According to the theory as developed by Helmholtz, the heat produced by the compression that has taken place is enough to account not only for the high temperature of the sun but also for the immense amount of heat energy that has been radiated. He estimated that a decrease in the diameter of the sun of 250 feet will produce the heat now radiated in one year. This rate of shrinkage is entirely too small to be detected in a thousand years. Helmholtz's theory is not accepted today as a complete explanation. However, there is no doubt that compression of the sun has played a very large part in the development of its heat.

229. Two important cases; free expansion. It is very important to distinguish between two cases of the expansion of gases.

1. If air or some other gas is allowed to escape from a tank where it is under pressure, the compressed gas in the tank does work in giving kinetic energy to the escaping gas. The gas left in the tank is cooled *because it does work*, and may become quite cold. The student should keep clearly in mind that the part of the gas which is cooled is the part that does the work. The escaping blast becomes cold because the compressed air from which it comes has become cold.

2. On the other hand, suppose that while the air is escaping, air is pumped into the tank at such a rate that the pressure in the tank is maintained constant. The air in the tank still does work

in giving to the escaping blast kinetic energy. But work is now being done on the air in the tank. Since the pressure remains constant, the air is compressed just as much as it expands. It can be proved that if a gas accurately obeyed Boyle's law, the work done on the gas by the force pump would be exactly equal to the work the gas does in giving kinetic energy to the escaping part. Since the total change in the energy* of the gas is zero, no temperature changes should be expected; at least, none from this cause. But since gases do not accurately obey Boyle's law, slight changes should be expected in the energy and hence in the temperature of the gas. In the case of some gases the escaping part is slightly cooled, but in others (notably hydrogen) there is a slight heating effect. In all cases the heating or cooling is very slight. Expansion of a gas under these circumstances, where the compressed gas in the tank is kept at a constant pressure, is called the *free expansion* of a gas. The slight heating or cooling on free expansion is sometimes called the *Joule-Thomson effect.*

It is important to remember the distinction between these two cases. Time and money have been wasted by inventors who, working on devices for supplying cooled air to buildings, railroad cars, etc., did not understand this difference. Some of them made preliminary experiments on the escape of air from tanks under the first condition, where the pressure was not kept constant, and expected to get the same cooling in a continuous process where the pressure in the tank was kept constant. The distinction is easy if one keeps in mind the fact that in the first case the gas does work without having any work done on it, and hence loses heat energy. In the second case there is approximately as much work done on the gas as it does.

230. Two specific heats of gases. In the case of gases two different specific heats are used:

1. *The specific heat at constant pressure is the quantity of heat required to change the temperature of 1 gram of the gas 1 degree centigrade, the pressure of the gas being kept constant.*

---

*When a gas expands, the potential energy of the molecules may change. This is neglected here. The effect, if any, is very small.

2. *The specific heat at constant volume is the quantity of heat required to change the temperature of 1 gram of the gas 1 degree centigrade, the volume of the gas being kept constant.*

The specific heat of any gas at *constant pressure* is always greater than the specific heat of that gas at *constant volume*. The student is expected to be able to state why this is true, after he has recalled two facts: (1) if the temperature is raised while the pressure is kept constant, the volume changes, and work is done by the gas as it expands; (2) when the volume is kept constant, the pressure increases, but no work is done by the gas because there is no motion of the walls of the inclosure. A gas does work only when the pressure it exerts actually moves something.

TABLE OF SPECIFIC HEATS OF COMMON GASES

	CONSTANT PRESSURE	CONSTANT VOLUME
Air	0.242	0.173
Hydrogen	3.40	2.40
Nitrogen	0.24	0.17
Oxygen	0.24	0.17

231. Heats of combustion. In many cases where chemical combinations take place, energy is liberated through the changes in the potential energies. For example, coal when burned liberates energy in the form of heat. Since energy is obtained from our food by the oxidation that takes place, the total amount of available energy in foods can be determined by measuring the amount of heat energy given off when they are burned (oxidized). Since energies of this kind are usually measured by the heat evolved, it is customary to express the energy content in heat units, in calories, or in British thermal units. The energy given off by 1 gram or 1 pound of the material is usually called the *heat of combustion*.*

---

*There is some confusion owing to the fact that in some of the literature *kilogram-calories* are called *calories*. For example, the quantities commonly used for expressing the energy in foods should be labeled *kilogram-calories* and not *calories*. The food ration of an adult ranges from 2000 to 4000 *kilogram-calories*, not calories.

HEATS OF COMBUSTION

(In Calories per Gram*)

Alcohol (ethyl)	7180	Coke	7000
Alcohol (wood)	5300	Gas (coal)	6000–11,000
Benzene	10,000	Gasoline	11,000
Coal (anthracite)	7600–8400	Hydrogen	34,000
Coal (bituminous)	6100–7800	Petroleum (crude)	11,000

232. Thermodynamics. The more advanced theoretical development of the relationships between *work* and *heat* is called *thermodynamics*. In a general sense it covers the principles relating not only to heat machinery, such as steam and gas engines, but also to other thermal devices, and to many chemical and electrical processes where temperature changes are involved. Some of the principles developed are now accepted as ranking among the fundamental laws of nature. Obviously it would be out of place to enter into details here. However, the *first* and *second laws of thermodynamics* will be explained. One important deduction applying to heat machinery will be given in Chapter XXIII.

233. The first law of thermodynamics. The first law is the formal statement of the facts given by equation (19) of section 224:

$$W = JH.$$

If the mechanical and the heat energy are measured in the *same* units, the first law of thermodynamics may be stated as follows: *In any transfer of other forms of energy into heat, or of heat into other forms of energy, the heat energy is equal to the amount of transformed energy.*

This law is merely a special case of the law of the conservation of energy.

234. The second law of thermodynamics. The second law of thermodynamics may be stated thus: *It is impossible to operate continuously any machinery by heat energy obtained from bodies the temperature of which is equal to or lower than that of surrounding bodies.*

---

* The heat of combustion in British thermal units per pound can be found by multiplying the values in the table by 1.80.

The oceans of the earth contain tremendous amounts of heat energy. If we could in some way extract that energy and put it to use, we should have available another great storehouse of energy. If we took heat out of the water, we should cool the water to a temperature below its surroundings. From the second law it follows that it is impossible to get useful work from the heat in the ocean. Useful mechanical energy can be obtained from heat energy only when the heat is taken from bodies which are warmer than their surroundings.

The first and second laws of thermodynamics cannot be proved by theoretical reasoning. They are the result of observation and experiment, and have stood the test of so many different types of experiments that we assume they are always true.

235. Dissipation and degradation of energy. There are two important tendencies in nature:

1. Vast amounts of energy are being transferred into heat energy, which is dissipated throughout space. This is true not only of energy supplied to inanimate machinery but also of that supplied to animal and human life.

2. The tendency is for all bodies to come to the same final temperature. When all bodies are at the same temperature, the heat energy they contain is, according to the second law of thermodynamics, no longer available for use. A hot body in cooling down to its surroundings gives up part of its heat energy to the neighboring bodies, and at the lower temperature this energy is no longer available for use. This process is called the *degradation of energy.*

These two tendencies show that while the total energy in the universe may remain constant, the part which is available for use by man is steadily becoming less. These conclusions, on account of their serious character, have attracted wide attention.

236. Summary. The old theory that heat was a substance has been abandoned, and heat is now believed to be a form of energy—kinetic energy of the molecules, atoms, and electrons.

The *calorie* and the *British thermal unit* are units for measuring energy.

The *mechanical equivalent of heat* is the number of ergs or joules in 1 calorie, or it is the number of foot-pounds in a British thermal unit.

In general, compression of a gas heats it and expansion of a gas cools it. In section 226 is given a reason for this, based on the kinetic theory of gases. It follows also from the law of conservation of energy. The *adiabatic* compression or expansion of a gas is a process so conducted that heat does not flow into or out of the gas during the compression or expansion.

There are two important cases of expansion: (1) Gas which is escaping from a container has kinetic energy. This energy comes from the thermal energy of the compressed gas in the container. Hence the compressed gas becomes colder. (2) If the pressure in the container is kept constant by some sort of pump, while the gas is escaping, the compressed gas in the container will do work as in (1), but work will be done on it in maintaining the pressure constant. If the gas obeys Boyle's law, the gas in the container will neither lose nor gain energy, and its temperature will not change. However, gases do not strictly obey Boyle's law, and there are slight temperature changes. In most cases the escaping gas is slightly cooled, but hydrogen at ordinary temperatures shows a slight heating effect. Expansion under these conditions is called *free expansion*, and the slight change in temperature that takes place is called the *Joule-Thomson effect.*

The *specific heat* of a gas at *constant pressure* is always greater than the specific heat of that gas at *constant volume.* (Why?)

The *heat of combustion* of a substance is the number of heat units 1 gram or 1 pound will give up when burned.

The *first law of thermodynamics* states that when mechanical energy is converted into heat energy, or vice versa, the amount of mechanical energy is equal to the amount of heat energy. If the two forms of energy are not measured in the same units, then, by equation (19),

$$W = JH,$$

where J is the *mechanical equivalent* of heat.

The *second law of thermodynamics* states that it is impossible to operate continuously any machinery by heat energy obtained from bodies the temperature of which is equal to or lower than that of surrounding bodies.

The amount of energy which is available for use by man is steadily decreasing through the *dissipation* and *degradation* of energy.

PROBLEMS

1. How does the experiment of Joule show that there must be internal friction or viscosity in water?

2. A motor with an output power of $\frac{1}{2}$ H.P. churns up 209 lb. of water for an hour, and as a result the temperature of the water rises 6.1° F. What value does this indicate for the mechanical equivalent of heat?

3. A bullet of 8 gm. mass has a velocity of 600 m./sec. To how many calories is this energy equivalent?

4. Find the distance a lead ball must fall if the impact warms the ball 1° C., assuming that half of the energy of fall goes to heat the ball. Neglect air resistance. (The specific heat of lead is 0.031 cal. per gram per degree.)

5. If a man eats 1000 kg.-cal. in 20 min., at what horse power is he receiving energy?

6. A man eats 3800 kg.-cal. each day and performs 600,000 ft.-lb. of useful work. Regarding him as a machine, compute his efficiency.

7. 1000 H.P. is required to run a railroad train at a constant speed on a level track. How many pounds of water could be raised from freezing point to boiling point per hour by this power?

8. If the earth receives from the sun 2 cal. of heat per minute for each square centimeter of the earth's surface, how many horse power does it receive per square meter?

9. A rifle bullet the weight of which is 0.0124 lb. has a velocity of 3000 ft./sec. To how many British thermal units is this equivalent?

10. A given type of coal has a heat of combustion of 7500 cal./gm. If 20 per cent of the energy could be converted into useful work, how much coal would be required for each horse-power hour of useful work?

11. $\frac{1}{2}$ H.P. is expended in the frictional losses in the shafting of a machine shop. How many British thermal units are generated per hour?

CHAPTER XX

THE TRANSFER OF HEAT

Methods of transfer of heat, 237. Conduction, 238. Rate of flow of heat; conductivity, 239. Convection, 240. Systems of heating, 241. Winds and ocean currents, 242. Radiation, 243. Heat radiation travels in straight lines, 244. Reflection, 245. Radiation at low temperatures; law of exchanges, 246. Apparent radiation of cold from ice, 247. The mean temperatures of the earth and other planets, 248. Cooling by radiation, 249. Good and poor radiators, 250. The Dewar flask; the thermos bottle, 251. Other illustrations, 252. Radiant energy not a homogeneous radiation, 253. Rate of cooling of hot bodies, 254. Summary, 255.

237. Methods of transfer of heat. Heat is transferred by three different methods: *conduction*, *convection*, and *radiation*. These, with numerous illustrations, will be explained in this chapter.

238. Conduction. If a silver spoon is placed in a cup of very hot coffee, the handle of the spoon becomes hot. When there is a fire in a stove, all metal parts of the stove get hot. In some way heat is *conducted* from the hotter parts of the spoon or the stove to the colder parts. This is often spoken of as a *flow* of heat from one part of a body to another. But the word *flow* may be misleading. Heat is not a fluid: it is the energy of agitation of the extremely small particles of which substances are constructed. When one end of a bar is heated, the molecules and electrons near that end are given greater vibratory motion, greater kinetic energy. In some way, probably by collisions, these particles are able to share their increased energy with their neighbors, the latter, in turn, sharing their increased kinetic energy with those beyond, and so on. Thus energy of agitation is propagated along the bar. The electrons, being very much smaller than the molecules and atoms and being able to move in the interstices between the molecules, are probably the most effective agents in transferring energy from one molecule to its neighbor.

Metals are very much better heat-conductors than other substances. Wood, for example, is a poor conductor. If you touch a piece of hot iron, and then a piece of wood at the same temperature, the iron feels hotter because it conducts heat to your hand at a higher rate. When the iron and wood are both cold, the iron will feel colder because it conducts the heat more rapidly away from your hand. A piece of paper wrapped tightly in one layer around a smooth metal rod can be held in a Bunsen gas flame without scorching it; but if the paper is wrapped around a wooden rod and placed in the flame, it will be quickly scorched. Why?

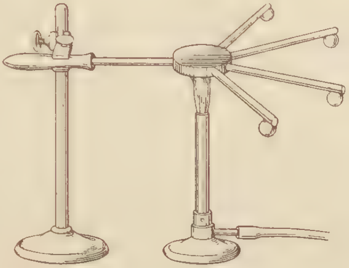


FIG. 155

Metals are not equally good conductors. A simple experiment shows the difference. Short, slender rods of the same length and diameter but made of different metals (copper, brass, iron, etc.) are mounted in a small disk of metal so that they extend out in different directions from the disk (Fig. 155). Small lead balls are fastened with wax near the outer ends of the rods. If a Bunsen flame is applied to the central disk, heat will be conducted from the disk out along the rods. The copper rod will be found to be a better conductor than the iron or brass rods, for the wax on its end will melt, and its lead ball will fall much sooner than the other balls. Of all metals silver is the best conductor, copper being next. Aluminum is among the best conductors, being better than brass and iron. (See table on page 254.)

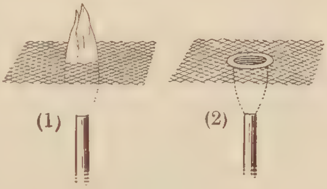


FIG. 156

An interesting example of the effect of conduction of heat is shown by a copper-wire gauze placed in a Bunsen flame. If the gauze is placed over the burner before the gas is lighted, and the gas is lighted above, the flame will stay above it (Fig. 156 (1)); or, if the gauze is brought down on the flame (Fig. 156 (2)), the flame will not strike through, although there may be much unburned gas flowing through the gauze. In the first case the gas can be extinguished by lifting up the gauze. The fact that a flame usually will not strike through a wire gauze is made use of in the construction of the lamps carried by miners. The flame of these lamps is surrounded with metal gauze, so that it may be carried into a mine where there are inflammable gases without igniting them. It is left to the student to explain why a flame will not strike through a relatively cold metallic gauze.

Snow is a very poor conductor of heat. When the ground is covered with a few inches of snow, it may not lose much heat even in severe winter weather. Under a heavy blanket of snow the ground often thaws from the heat of the deeper layers of the ground. Gases are all poor conductors; hence loosely woven fabrics, which contain a great deal of air, are poor conductors. In the building of houses and refrigerators, and in other cases where poorly conducting walls are wanted, air cavities are often left. The efficiency of refrigerators can usually be improved by covering the outside with layers of paper; for paper is a poor conductor, and the walls of refrigerators are often not built as well as they should be. Liquids are relatively poor conductors, although they are better than gases.

Fireless cookers are vessels surrounded by poor conductors. Cooking usually requires that the material be raised to a high temperature and maintained at that temperature for a certain period. The process does not require energy other than that necessary to raise the temperature. If, after the food is thoroughly heated, it is placed in a thermally isolated vessel (in a fireless cooker, for example), it will stay at a high temperature a long time and will continue to cook. If the insulation were perfect, it would never cool.

Meats, especially those containing much fat, are not good conductors of heat. The length of time required for roasting meat is due to its poor conductivity. In good cuts it is necessary only to raise the temperature of the interior to a certain point; it is not necessary to maintain it at that temperature. Metal rods, or skewers, run through the meat should help. Why?

In the cold-pack method of canning, the material to be canned is placed in glass jars, and then heated by immersing these jars in hot water or steam. Glass is a poor conductor; the material in the jar is usually a poor conductor also, and it is often so tightly packed that currents in the liquid are prevented. The result is that the time required to raise the temperature of the interior may be considerable. Cake is often baked in pans with a central tube running up through the cake. This helps to bring the interior of the cake more quickly to the desired temperature. The advantages in the use of copper and aluminum, which are good conductors of heat, in cooking-utensils should be obvious. However, there are cases where it may be a disadvantage to use too good a conductor. For example, when cooking over a small gas burner, heat is conducted not only into the interior of the vessel but out to the sides and top, where it is dissipated into the room.

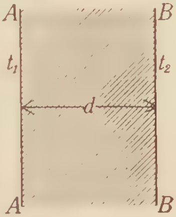


FIG. 157

239. Rate of flow of heat; conductivity. In many instances it is necessary to make accurate estimates of the rate of flow of heat through conductors. In order to understand the method used, certain quantitative concepts must be learned. In Fig. 157 is shown a section of a conductor. The two faces *AA* and *BB* are flat, and perpendicular to the paper. Let one face be maintained at a temperature t_1, and the other at a lower temperature, t_2. It is found by experiment that the rate of flow of heat is (1) proportional to the difference of temperature, $t_1 - t_2$, of the two faces; (2) inversely proportional to the thickness d of the slab; (3) proportional to the area of the faces *AA* and *BB*. These facts may be stated by the following equation:

$$\text{Rate of flow of heat in calories per second} = k\frac{t_1 - t_2}{d} \times \text{area.} \quad (20)$$

The student should see that this includes all three statements just given. This equation applies only to the case where the temperatures are constant and the flow has become steady.

For any one material k is a constant, which is called the *coefficient of thermal conductivity*, or usually the *thermal conductivity*.

When the *thermal conductivity*, k, of any substance is known, the rate of flow of heat through that substance can be computed if the dimensions and the difference of temperature are known.

THERMAL CONDUCTIVITIES

(In Calories per Second per Centimeter per Degree Centigrade)

Aluminum	0.50	Iron (wrought)	0.14
Brass (70 per cent of copper, 30 per cent of tin)	0.26	Iron (steel)	0.11
Copper	0.91	Lead	0.09
Glass (window)	0.0025	Mercury	0.019
Gold	0.7	Nickel	0.14
Iron (cast)	0.15	Sand (white, dry)	0.0009
Iron (pure)	0.16	Silver	0.97
		Tin	0.15

Obviously, if equation (20) gives the rate of flow of heat, or the flow for one second, the total flow is given by the equation

$$\text{Flow of heat in calories} = k\frac{t_1 - t_2}{d} \times \text{area} \times \text{time in seconds.} \quad (21)$$

The fraction $\frac{t_1 - t_2}{d}$ is often called the *temperature gradient*. If d is measured in centimeters, the temperature gradient is the fall in temperature per centimeter. It is analogous to the gradient, or grade, of a railroad track or roadway. Equation (21) shows that the flow of heat is proportional to the gradient.

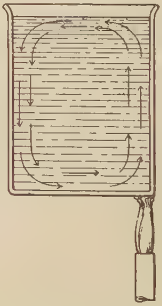


FIG. 158

240. Convection. When a flame is applied to a vessel of water, *convection* currents of water are set up (the direction is indicated by the arrows in Fig. 158). When the temperature of the water immediately over the flame is raised, the water expands and becomes less dense than the surrounding, colder water. The buoyant effect of the colder water causes the warmer water to rise, the colder water taking its place. Then the colder water, becoming heated, rises in turn, and thus a continuous circulation or stirring of the water takes place. On account of the motion of the water there is a tendency for it to come rather quickly to a nearly uniform temperature. This method, called *heating by convection*, is the usual one by which liquids and gases are heated. It is the method used in our hot-air furnaces.

Why is a radiator of a house-heating system more effective when placed near the floor than near the ceiling?

Should ice be placed near the top of the refrigerator or near the bottom? Why?

241. Systems of heating. A thorough knowledge of the simple principle of convection is of great practical importance because of the common use of this principle in our heating systems. In Fig. 159 are shown the more essential parts of a hot-water heating system. There is a continual circulation of the water by convection, the hot water rising from the heater and flowing through the radiators and back to the heater.

The heating system shown in Fig. 159 is called the *direct-indirect-radiation* system, because heat is supplied directly by the radiators and also by fresh air which is heated as it passes through the radiator. (The use of the words *radiators* and *radiation* in this connection is somewhat misleading. This will be obvious after the student has read the sections on radiation.) The *direct* system, in which there is no provision for ventilation, is the one commonly used in steam and hot-water heating.

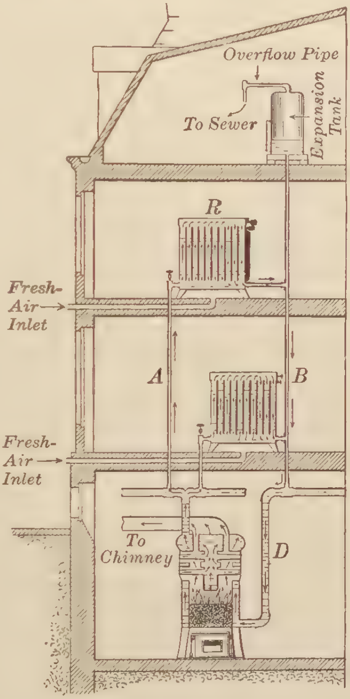


FIG. 159. Hot-water heating

In the *indirect* system air flows over steam pipes in the basement, the hot air being distributed by a method similar to that used with hot-air furnaces. Often in the indirect system the air is forced through by large fans, which insure adequate circulation and ventilation.

The subject of ventilation is a very important one, far more important than usually believed. While the underlying principles are simple, it is so important a matter that only experts should attempt to design or work out the proper ventilation of public buildings. A fact especially worth mentioning is that in order to cool a room quickly in the evening after a hot day it is of great advantage to have openings in or near the ceiling. Casement windows are far more effective than the ordinary windows, which are usually open only at the bottom. On the other hand, in winter weather it is usually a great disadvantage to have places in the ceiling for the air to leave the room, as is a common practice in some churches and public halls. In that case the fresh air entering from the registers, being the warmest air in the room, rises at once and flows out through the ventilators in the ceiling.

242. Winds and ocean currents. The chief causes of winds are the convection currents caused by the unequal heating of the surface of the earth. The air near the equator rises and flows toward the poles, while the surface layers of air tend to flow toward the equator. These currents are greatly modified by the rotation of the earth, by local disturbances due to unequal heating, by mountain ranges, and the like. If one has watched the currents of water in a large vessel over a hot fire and noticed the turbulent conditions in the water, he can better form a mental picture of the turbulent conditions existing on a vast scale in the earth's atmosphere.

In the daytime the land is, in general, warmer than the surface of the ocean. The large specific heat of water and the fact that part of the sun's heat penetrates more deeply into it prevent as large a rise of temperature of the surface of the water as of the land. Hence the air lying over the land becomes hotter than that over the water, and a convection current is set up, with the surface wind blowing from the water toward the land—the so-called sea breeze. At night, for a somewhat similar reason, the air flows in the opposite direction, giving the land breeze. When either the land or the sea breeze coincides in direction with the prevailing wind, it will be very strong. When the prevailing wind is from the water, the day wind will be strong; at night, on account of the two opposing effects, the breeze may be very weak.

Mountain and valley winds are quite noticeable. In the daytime the wind blows up the mountain; but at night the surface of the mountain cools off quickly and cools the air, causing a downward flow. When this downward breeze converges from two or more mountain sides into a narrow valley, the breeze may become strong. In rolling country, convection currents are more readily produced, with the result that the summer nights are more comfortable than in flat prairie country. (See also section 295.)

The currents of water in the ocean are often well defined, and are explained by the simple principles of convection, the interference of masses of land, the effects due to the earth's rotation, and the action of the prevailing winds.

243. Radiation. The process of transfer of energy known as *radiation* is very different in character from *conduction* or *convection.* A few examples will assist in making clear to the student just what is meant by the term *radiation.* By some process energy traverses the space between the sun and the earth, although this space is undoubtedly a more perfect vacuum than any that we can make in our laboratories. This energy when absorbed becomes heat energy. The heated filament of an incandescent lamp, even when there is no air or gas in the glass bulb, emits energy which produces heat where this energy is absorbed. One can readily detect this by holding the hand or the face near such a lamp. The same effect is observed near any heated object, such as a steam pipe, a stove, or the glowing coals in a fireplace. In most of these cases the student can easily satisfy himself that the heat is not brought to his hand or face by conduction or by convection currents. The term *radiation*, when used in a general sense, is a process of transferring energy from a body across space.

It is now generally accepted that that form of propagation of energy which we call heat radiation is a wave-motion. The reasons for this belief will be given in the chapters on light. In some way the rapid vibrations of the electrons and atoms in a hot body are able to produce waves which travel across space to another body and there stimulate vibrations of the electrons and atoms, thus producing heat. Note that the wave-motion is not heat, but produces heat when absorbed.

Many of the facts concerning heat radiation given in the following sections can be easily demonstrated by the use of a simple air thermometer. It consists of a glass bulb, four or five centimeters in diameter, with a long narrow stem. It should be inverted with the end of the stem below the surface of some colored water. The bulb must be blackened. If the bulb is slightly heated, the inclosed air will expand, part bubbling out through the water. On cooling, the colored liquid will rise part way up the stem, where it will be plainly visible. A more accurate instrument is the *thermopile*, which utilizes the thermoelectric phenomena referred to in section 190.

244. Heat radiation travels in straight lines. It is very easy to show that heat radiation travels in straight lines, for any kind of obstacle placed between a hot body and the blackened bulb of an air thermometer will cut off the heating effect. A fine illustration of this effect is the fact that when the moon comes between the earth and the sun, producing a solar eclipse, heat as well as light is cut off.

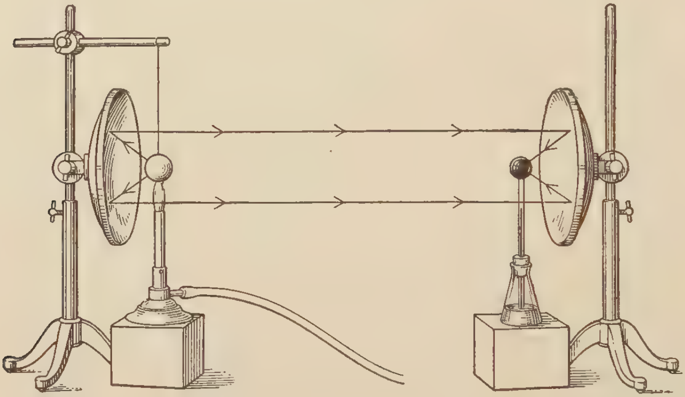


FIG. 160

245. Reflection. With a pair of large spherical metal mirrors the radiation from an iron ball hanging in a Bunsen flame can be reflected and focused on the bulb of an air thermometer. The effect is quite decided even when the mirrors are four or five meters apart (Fig. 160).

246. Radiation at low temperatures; law of exchanges. It has been shown that radiation takes place from all bodies, no matter what their temperature may be. Even ice radiates some energy. However, the higher the temperature of a body, the greater the rate at which it radiates energy. If a cold body is brought into a warm room, it will radiate energy to the other objects in the room; but these other bodies, being warmer, will radiate more energy to it than it gives out. The temperature of the cold body will gradually rise until it comes to the same temperature as the other bodies in the room, when it then will receive radiant energy at exactly the same rate as it emits it. The principle is known as *Prevost's law of exchanges.*

247. Apparent radiation of cold from ice. If two concave reflectors are set facing each other, and a block of ice is placed at the focus of one mirror and the blackened bulb of an air thermometer at the focus of the other, the temperature of the air thermometer will fall. At first sight this seems to be due to a radiation of *cold* from the ice. The true explanation is as follows: The ice radiates some energy, which is reflected over to the bulb of the air thermometer. The bulb of the air thermometer radiates some energy which is reflected over to the ice. There is thus an exchange going on. But the bulb of the air thermometer, being at a higher temperature, radiates more energy to the ice than it receives from the ice. It is therefore cooled by its own radiation.

248. The mean temperatures of the earth and other planets. The earth receives from the sun radiant energy, which tends to raise the temperature of the earth. The mean temperature of the earth is determined by the fact that this temperature must be such that the earth will radiate energy at the same average rate at which it receives energy from the sun. Since the rate at which energy is received from the sun is known (at least approximately), it is possible, by using the radiation law of section 254, to compute what the temperature of the earth must be in order to radiate energy at this rate. This computation gives a result near the observed average temperature of the earth.

The planet Venus is much nearer the sun, and hence receives radiant energy at a greater rate than we do. It must therefore be at a higher temperature than the earth in order to radiate energy at the same rate as it receives it. Estimates of its mean temperature based on this kind of reasoning indicate that the mean temperature of Venus is about 80° C., or about 175° F. Mars, on the other hand, is farther away from the sun than the earth. It receives less radiant energy; and in order that it shall radiate energy at the same average rate as it receives energy, its average temperature must be lower than that of the earth. A simple estimate indicates that its mean temperature is about the freezing temperature of mercury; that is, about $-40°$ C. Estimates of this kind are not very accurate, for the type of atmosphere and the amount and kind of cloudiness greatly modify the results.

249. Cooling by radiation. On clear nights objects lying out in the open radiate energy into the space above them and do not receive much energy in return. They therefore cool off rapidly and fall to a temperature lower than that of the air. This is especially noticeable if they are not good conductors, and hence do not receive heat by conduction from the ground below. Straw, boards, and other nonconductors usually get colder than the air on clear nights. Such things as rocks and cement walks are kept warm by conduction of heat from the ground. This effect of the cooling due to radiation is very obvious in the spring and fall when frosts are likely to occur: frost usually forms only on objects which get colder than the air. On clear nights when the temperature of the air goes no lower than 35° F., water in shallow pans placed on boards or on straw away from buildings and trees will usually freeze. On cloudy nights the clouds radiate energy to the earth and thus prevent the temperature of plants and other objects from falling below air temperature.

250. Good and poor radiators. Different kinds of surfaces show very different radiating properties. If one side of a polished vessel is coated with lampblack, the blackened side, when the vessel is filled with boiling water, will radiate more energy than the polished side. This can easily be shown by bringing the vessel near the blackened bulb of an air thermometer and turning first one side and then the other toward the thermometer. If two vessels which are similar, except that one has its surface blackened and the other is polished, are filled with boiling water and allowed to cool, it will be observed that the temperature of the blackened one falls more rapidly than that of the other. A blackened surface is a better radiator than a polished one.

It is a general principle that *a good radiator is also a good absorber of radiant energy.* This can be easily shown by theoretical considerations. Imagine that a body *A* which has a good radiating surface is placed inside some inclosure where there are other bodies. A closed room would do. This body *A* will radiate energy to the room and will also absorb energy which is radiated by the other bodies. It will finally come to the same temperature as the room, and will then absorb energy at the *same rate* as it radiates it. From this one can conclude that if it is a good radiator (that is, if it radiates energy rapidly), it must also be a good absorber; for otherwise it would get colder than the room. Similar reasoning will show that *if the surface of a body is a poor radiator, it must also be a poor absorber.*

Poor absorbers are those that reflect or transmit radiant energy; hence *good reflectors must be poor absorbers.* It must also be true *that good reflectors are poor radiators*; for instance, coffeepots and teapots which have bright, reflecting surfaces are poor radiators.

251. The Dewar flask; the thermos bottle. The Dewar flask is ingeniously constructed to prevent heat from either entering or leaving it. It is a double-walled glass bottle with a vacuum between the two walls. This vacuum prevents any flow of heat in or out of the interior by *conduction* or by *convection*, except for a slight amount of conduction where the walls join at the neck. The walls of the flask are silver-plated on the side next to the vacuum. These reflecting surfaces are both poor absorbers and poor radiators, thus preventing to a large extent the radiation of heat through the vacuum. Flasks of this kind were first designed by Professor Dewar as containers for liquid air. They are sold for domestic purposes under various names, the best-known being "thermos." It should now be clear why one of these flasks will keep hot fluids such as coffee, or cold things such as ice cream, for many hours without much change in temperature.

252. Other illustrations. It is often noticed that dirty snow melts rapidly in strong sunlight, while clean snow close by is scarcely affected. Why is this?

The bark of most trees is a good absorber of radiant energy. A small thermometer slipped under the bark shows that on the south side of a tree, on a sunshiny day in winter, the temperature gets considerably higher than the air temperature, and higher than the temperature under the bark on the north side of the tree. In certain types of trees trouble often develops on the south side of the trunk. This is frequently explained by the fact that in winter time the south side has abnormal changes in temperature between day and night. It is claimed that these troubles can often be eliminated by protecting the trunk of the tree from sunlight.

In many parts of the United States, peach buds will start growing in the winter, and later will be killed by low temperatures. The surface of the twigs is a good absorber of radiant energy from the sun; hence on clear days the temperature of the twigs will often rise considerably above air temperature. This causes premature growth. Attempts are under way to develop a peach tree with twigs which do not have such good absorbing surfaces.

Air is not entirely transparent to radiant energy. It absorbs some of the radiation from the sun. In the daytime the rise in temperature of the air is in large part due to the absorption of energy radiated from the sun and of that radiated from the ground. After sunset, if the sky is clear, the air usually cools rapidly on account of the radiation from the air into space.

253. Radiant energy not a homogeneous radiation. The radiant energy from any white-hot body consists of waves the length of which varies through a very wide range. The waves which have wave-lengths between 0.00004 and 0.00007 centimeters (approximately) are able to produce the effect on our eyes which we call light. Some of the waves produce very marked chemical actions, as is shown in photography and in the effect of light on growing plants. Longer waves produce no chemical effects whatsoever: the only effect they have is to develop heat. For that reason they have been called *heat-waves*. But this term is very misleading. *Waves of all lengths will produce heat when absorbed.*

The radiation from a ball of iron not quite at red heat consists of waves which are too long to affect the retina of the eye. As the temperature of the ball is raised it will emit, in addition to the long waves, shorter and shorter ones. It first appears red (the longest visible waves give the sensation we call red) and later white (when it is giving out all colors).

Since different sources of radiant heat may emit different radiations, we should expect to find differences due to the type of radiator used. For example, it is found that a piece of plate glass is opaque to the radiation from a steam or hot-water radiator, but not to that of an incandescent lamp, an arc light, or sunlight. The reason for this is that while glass is quite transparent over the range of wave-lengths that can produce the sensation of light, it is not transparent for very long waves.

There are numerous cases where a body is a good absorber, and hence a good radiator, for some radiations, and a poor absorber and radiator for radiations of a different wave-length. White lead is a good example. It is a poor absorber for light waves, but it is a very good absorber for longer waves. One cannot tell from the color of a body whether it is a good or a poor absorber for all wave-lengths.

A large percentage of the radiant energy from the sun consists of waves that produce the sensation of light; but in our artificial sources of light only from 1 to 10 per cent of the radiant energy can produce a sensation of light, the remainder being chiefly in the longer waves.

254. Rate of cooling of hot bodies. The rate of emission of radiant energy is accurately known in the case of a "black" body; that is, one which absorbs 100 per cent of the energy falling on it. If E represents the number of ergs of energy radiated per second from each square centimeter of a black body, then

$$E = cT^4, \tag{22}$$

where T is the absolute temperature of the body and c is a constant equal to about 5.7×10^{-5}. This is known as *Stefan's law.* It was first suggested as an empirical law, because it agreed with experimental results; but later it was deduced theoretically, by using the fundamental principles of thermodynamics.

If a black body is surrounded by other bodies which are at a temperature of T_1, then it will receive radiant energy from them, and its net loss by radiation per second, or the rate of cooling, is

$$c(T^4 - T_1{}^4),$$

where T is the temperature of the black body. While this law is strictly true only for black bodies, it is approximately true for most substances.

Newton's law of cooling includes losses by convection currents in the air, as well as radiation losses. His law states that *the rate of loss of heat from any body to its surroundings is proportional to the difference in temperature between that body and its surroundings.* This law is approximately true when the difference of temperature is not large.

255. Summary. There are three different ways in which heat is transferred from one place to another: *conduction*, *convection*, and *radiation.*

Raising the temperature of a body increases the kinetic energy of the electrons, atoms, and molecules of which the body is composed. Probably through collision there is a tendency for these small particles to share their kinetic energy with their neighbors. Hence, if one end of a bar is heated, energy will be carried from that end to other parts of the bar. This process is called *conduction.* By equation (20)

$$\text{Rate of conduction of heat in calories per second} = k\,\frac{t_1 - t_2}{d} \times \text{area},$$

where k, the *thermal conductivity*, is a constant for any given substance.

Convection in gases and liquids is the process in which currents are set up by the changes in density which occur when the temperature changes. This is the usual method for heating liquids and gases.

Radiation of heat is a process in which energy is propagated in the form of waves. When these waves are absorbed by a body, they produce the energy of vibratory motion which we have called heat. This radiation travels in straight lines and can be reflected. All bodies radiate energy, whatever their temperature may be. Hot bodies radiate it at a greater rate than colder ones (see Stefan's law, equation (22)). It is important to distinguish between *good* and *poor* radiators. Poor absorbers are always poor radiators; good absorbers are always good radiators.

The waves which transfer radiant energy are not all alike, for they vary in wave-length through a wide range. Waves of all lengths transmit energy and, when absorbed, produce heat. But the greater part of the energy radiated by hot bodies consists of waves longer than light-waves.

A body which will absorb 100 per cent of the radiant energy falling on it is a perfect absorber. Such a body is often called a *black* body. It will radiate, for the same temperature, more energy than one which is not a perfect absorber. *Stefan's law* states that *the rate of radiation of energy from a black body is proportional to the fourth power of its absolute temperature.*

Newton's law of cooling is an approximate one, applying only when the difference in temperature between the body and its surroundings is small. This law states that rate of loss of heat is proportional to the difference in temperature between the warm body and its surroundings.

PROBLEMS

1. In the summer, when the air inside a house is cooler than out of doors, in what direction will the air in a chimney tend to move?

2. In cold weather there is a tendency for air to leak in around the windows of the first floor of a heated house and to leak out around the windows upstairs. Explain the cause.

3. It has been found that a tent which has been painted on the inside with an aluminum paint is cooler on a sunny day than one which has not been so painted. Why?

4. How many calories will be conducted per second through each square centimeter of surface of an aluminum plate 3 mm. thick if the difference in temperature of the two faces is 20° C.? (See the table on page 254.)

5. How many calories will be conducted per second through 1 sq. m. of window glass 2 mm. thick if the difference in temperature of the two faces is 10° C.? (See table, p. 254.)

6. The flow of heat through an aluminum cooking-vessel is 50 cal. per second per square centimeter. If the aluminum is 2 mm. thick, what is the temperature difference between the two faces of the metal? (See the table on page 254.)

7. Make a table showing the relative conductivity of silver, copper, aluminum, and cast iron. Take silver as 100, and compute the corresponding numbers for the other three. (See table, p. 254.)

8. A black body is maintained at a temperature of 1500° C. How many ergs of energy will be radiated per second from each square centimeter?

9. If the absolute temperature of a black body is doubled, by what factor will the amount of radiated energy be increased?

10. How much food, measured in kilogram-calories, must a man eat per day to compensate for a loss by radiation and conduction at the rate of 75 watts (75 joules per second)?

CHAPTER XXI

CHANGE OF STATE

Molecular states, 256. Fusion, 257. Changes in internal energy on fusion, 258. Heat of fusion, 259. Experimental method for determining the heat of fusion of ice, 260. Some effects due to heat of fusion of ice, 261. Changes in volume on solidification, 262. Effect of pressure on the melting point, 263. Undercooling, 264. Freezing points of solutions, 265. Evaporation; the kinetic theory, 266. Saturated vapors, 267. Nonsaturated vapors, 268. Dependence of pressure of saturated vapor on temperature, 269. Boiling, 270. Effect of pressure on the boiling point, 271. Why the boiling point depends on pressure, 272. Pressure of saturated vapors of different liquids, 273. The boiling points of different liquids, 274. The pressure exerted by a vapor is independent of the presence of other gases or vapors, 275. Distillation, 276. Fractional distillation, 277. The boiling points of solutions, 278. Heat of vaporization, 279. Evaporation a cooling process, 280. Freezing water by its evaporation, 281. Carbon-dioxide snow, 282. Critical temperatures, 283. The spheroidal state, 284. Ice and refrigerating machines, 285. Liquefaction of air, 286. Summary, 287.

256. Molecular states. When ice melts and changes into water, there is a change in the state of the molecules. The change of water into vapor is another type of change in state. The phenomena connected with these two changes and the reverse processes will be discussed at length in this chapter. The student will find that there are many interesting facts associated with them, and that a relatively small number of laws or principles is involved.

257. Fusion. One of the conspicuous facts connected with the change of a solid to a liquid is that for many substances this change takes place at a very definite temperature. For example, in the case of ice and water the temperature, 0° C., or 32° F., is so definite that it is used as a standard. While these temperatures are usually referred to as the melting-point temperatures, they are also the freezing, or solidifying, temperatures. For example, ice melts at the same temperature at which pure water freezes.

TABLE OF MELTING POINTS AT A PRESSURE OF 76 CENTIMETERS OF MERCURY

	Degrees Centigrade		Degrees Centigrade
Aluminum	658.7	Nickel	1452
Cadmium	321	Platinum	1755
Copper	1083	Silver	960.5
Gold	1063	Tin	231.9
Lead	327	Tungsten	3000
Mercury	−38.7	Zinc	419.4

In the case of many substances the change from a solid to a liquid does not take place at a definite temperature; for example, such substances as paraffin wax, sealing-wax, and butter gradually change from a solid to a liquid. The same thing is true of iron. One reason that iron can be welded is that it does not have a definite melting temperature, but softens and becomes plastic. If butter had a definite melting point, it would be difficult to use it as we now do.

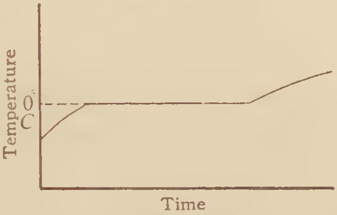


FIG. 161. Change in temperature of snow

258. Changes in internal energy on fusion. It is found by experiment that heat energy must be added to all substances to change them from solid to liquid. In the case of ice a very considerable amount of energy must be added to change it to water, even when there is no change in its temperature; hence a mass of 100 grams of water at 0° C. contains more internal energy than 100 grams of ice at 0° C.

If heat is supplied to snow or crushed ice at a uniform rate, and the temperature, as indicated by a thermometer, is observed at regular intervals, results similar to those shown in Fig. 161 will be obtained. As shown, the temperature of the snow rises quickly to 0°, when it begins to melt. The mixture of snow and water (which should be kept well stirred) will stay at that temperature, although heat is being supplied, until the snow is practically all melted. After that the temperature will rise. This experiment shows clearly that heat energy is needed to change ice to water;

for during the time the ice was melting it was receiving heat from outside, and yet the temperature did not rise.

If water is taken into a cold place where the temperature is considerably lower than 0° C., and its temperature observed at regular intervals, results similar to those shown in Fig. 162 will be obtained. The flat part of the curve shows a constant temperature during the change of state from water to ice. During the time water is *changing into ice it gives up heat energy.*

Experiments similar to those represented by Figs. 161 and 162 are often performed in order to determine the melting, or fusing, points of substances. When a substance has a definite melting point, the flat part of the curve is usually so clearly shown that there is little trouble in determining the temperature corresponding to it.

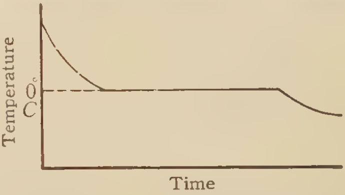


FIG. 162. Cooling and freezing of water

259. Heat of fusion. It has just been pointed out that when ice changes into water, heat energy must be added; and that when water changes into ice, heat energy is given up. Careful experiments have shown that the amount of heat that must be added to melt 1 gram of ice at atmospheric pressure is always the *same in amount.* This quantity is called the *heat of fusion* of ice.

The heat of fusion of any substance is equal to the number of calories of heat required to melt 1 gram of the substance, its temperature remaining unchanged.

260. Experimental method for determining the heat of fusion of ice. A simple method for determining the heat of fusion of ice is based on the method of mixtures. A description of an experiment, with numerical results, will assist in a clearer understanding.

A copper vessel of mass 100 grams and specific heat 0.09 contained 400 grams of water at 25.2° C. A "dry" piece of ice (that is, one without water clinging to it) at 0° C. was dropped into the vessel, and after the ice was all melted the final temperature was 13.5° C. The mass of the vessel and its contents was found to be 550.9 grams. The increase in the mass gives the mass of ice added, which is found to be 50.9 grams.

The number of calories given up by the copper vessel and the 400 grams of water in cooling from 25.2° C. to 13.5°C. was

$$100 \times .09 \times (25.2 - 13.5) + 400 \times (25.2 - 13.5) = 4785.3 \text{ cal.}$$

If the heat of fusion of the ice is represented by H, the number of calories required to melt 50.9 grams of ice and to raise its temperature to 13.5° C. is $50.9\,H + (50.9 \times 13.5)$.

Equating the heat given up by the vessel and water to the heat received by the ice, we have

$$4785.3 = 50.9\,H + (50.9 \times 13.5);$$

whence $$H = 80.5 \text{ calories per gram.}$$

This value is a little too high. More careful experiments, with corrections for heat received or given up to the room and for other experimental errors, give, as the heat of fusion of ice,

$$H = 79.8 \text{ calories per gram.}$$

HEATS OF FUSION

(In Calories per Gram)

Aluminum	80	Naphthalene	36
Glycerin	42	Sulfur	10
Lead	5	Tin	15
Mercury	3	Water	79.8

261. Some effects due to heat of fusion of ice. The relatively large heat of fusion of ice (approximately 80 calories per gram) plays an important part in the use of ice in refrigerators. Ten kilos (about 22 pounds) of ice will absorb, in melting, about 800,000 calories. This is enough to lower the temperature of 50 kilos of water 16° C. If the heat of fusion of ice were as small as that of lead or mercury, the use of ice in refrigerators would be of little service.

When water changes into ice, it gives up the same number of calories as it absorbs when changing from ice to water. The reason why large masses of water tend to keep the air temperature from going much below the freezing point should be obvious. When the ground is very wet, we are not as likely to have a hard freeze as when the ground is dry, other conditions being the same. If the heat of fusion of ice were a small quantity, ponds would more frequently freeze to the bottom in winter.

262. Changes in volume on solidification. When water freezes, it expands. Two evidences of this are well known: first, the smaller density of ice, as indicated by its floating on water; second, the disastrous effects of the freezing of water pipes. The specific gravity of ice at 0° C. is 0.92.

The expansion of water on freezing disintegrates rocks and loosens up the soil, an effect which is very beneficial for plant growth. In a stony field the tendency of freezing and thawing to bring rocks to the surface is very noticeable. The freezing of the ground underneath raises the rocks; when the ground thaws, the soil is usually washed under the rocks. The raising of stakes and posts which have not been set deeply in the ground is quite common. The roots of winter wheat, blue grass, and other plants with roots which lie near the surface are raised out of the ground by frequent freezing and thawing. It is well known that the footings or foundations of buildings and monuments must be placed below the line of possible freezing. Cement steps and porches are often broken away from buildings on account of the expansion of the ground underneath on freezing.

Waxes, gelatin, and many metals—in fact, most substances—contract on solidification; water is an exception to the general rule. Gold and silver coins cannot be made by casting; they must be stamped, for the contraction on solidification prevents them from accurately fitting their molds. The manufacture, by casting, of the type used in printing requires a metal which expands on solidification.

263. Effect of pressure on the melting point. An increase of pressure lowers the temperature at which water freezes or ice melts. The effect is very easily remembered; for when water freezes, it expands, and pressure tends to prevent this.

It is a general rule that *an increase of pressure lowers the temperature of the melting points of those substances that expand on* *solidifying and raises the melting points of those that contract on solidifying.* The changes in the melting points produced by pressure are not large. For water there is a decrease of 0.0072° C. in the melting point for each additional atmosphere of pressure.

It is commonly known that snowballs are easily made only when the snow is at the melting point and is wet; that is, contains some water. When a mass of snow is pressed in the hands, the pressure lowers the melting point. But this snow, which was at 0° C., cannot stay at a temperature higher than its melting point. Hence some of the snow melts, a process which requires heat and which causes the temperature to fall to the new melting point. The result is a mixture of snow and water at a temperature slightly below 0° C. When the pressure is removed, the melting point changes to 0° C. But the temperature of the mixture is still below 0° C., and water in contact with snow or ice cannot remain at a temperature below its freezing point. Hence some of the water freezes, liberating heat until the temperature of the mixture is raised to 0° C. This partial freezing of the water in the mixture tends to bind the whole firmly together. Similarly, when two pieces of ice at 0° C. are pressed tightly together, they will freeze together when the pressure is relieved. The ice left in the tracks of a wagon in snow near melting temperature is another example.

This process of melting under pressure and freezing again after the pressure is removed is called *regelation.*

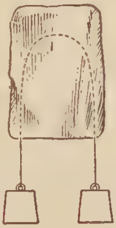


FIG. 163

It is worth while to explain fully a well-known experiment. If a fine wire which has heavy weights attached to its ends is hung over a block of ice, it will slowly melt its way through the ice, but its track will be left full of ice (Fig. 163). Ice is continually melted below the wire, and the water frozen above the wire. The explanation is as follows: The pressure produced by the wire will *lower* the melting point of the ice immediately under the wire. Since ice cannot exist at a temperature higher than its melting point, the temperature of the ice under the wire will be below 0° C. (Imagine what would happen if ice were at a temperature above its melting point. Some of the ice would immediately melt, the necessary heat being given up by the surrounding ice until it cooled down to the melting point.) Above the wire the temperature of the water and neighboring ice will be at the melting point, 0° C. Since the temperature on the top of the wire is higher than it is underneath, *heat will be conducted* across the wire. Heat will be continually taken away from the region above the wire and added to that below. Adding heat to the ice below the wire means that the ice there melts; taking heat away from above the wire, from the mixture of ice and water at 0° C., means that the water above the wire freezes. Since the amount of heat added below the wire is equal to that taken away from the region above, it follows that the amount of ice melted should be equal to the amount of water frozen. Attention may be called to the fact that pressure* alone does not melt ice,—that the melting requires heat (about 80 calories for each gram of ice melted).

The rate at which the wire will melt through the block of ice will depend on the diameter of the wire and on the weights used; for these determine the amount of pressure and hence the difference of temperature between the top and bottom of the wire in the ice. The rate depends also on how good a conductor of heat the material of the wire is. A copper wire will melt through faster than an iron wire of the same size with the same amount of suspended weight.

264. Undercooling. When pure water is cooled below 0° C. in a very clean glass vessel, it will often go below 0° C., sometimes to – 10° C., without freezing. This phenomenon is called *undercooling.* Ice has a crystalline structure in which the molecules have a very definite symmetrical arrangement. In the development and growth of a crystal there must be a nucleus on which it can start growing. If a small bit of ice is dropped into undercooled water, ice forms very rapidly, and the temperature rises to 0° C. The rise in temperature is produced by the heat given up by the freezing water. In laboratories the undercooling of water is not uncommon when distilled water and clean glass vessels are used. It will sometimes be observed in winter weather when water has been left standing in some clean vessel in an unheated room. There are reasons for believing that the minute drops of water which form clouds are often carried without freezing to high altitudes where the temperature is below 0° C.

---

* When there is motion as the result of pressure, work is done, and this work may be converted into heat. But simple computation, using the numerical values of section 224, shows that the work done is rarely sufficient to melt much ice. It is interesting to compute the amount of ice that would be melted by the work done when the weights of Fig. 163 fell a distance equal to the thickness of the block of ice.

Fig. 164 shows an example of the undercooling of acetamide, a substance which usually undercools before solidifying. The crosses in this figure give to scale the actual readings. As seen from the curve, the acetamide cooled down to a temperature of about 69° C., and then rose abruptly to its melting point, 76.5° C. In the figure the curve near 76° is rounded off. This is due to the fact that the thermometer was sluggish and did not instantaneously follow the changes in temperature.

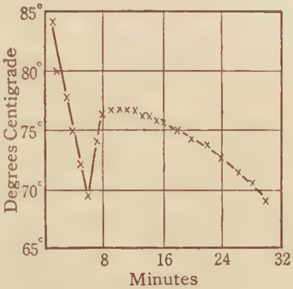


FIG. 164. Undercooling of acetamide

265. Freezing points of solutions. When a salt is dissolved in water, the freezing point of the solution is lower than that of water. For example, a saturated solution of common salt in water freezes at about 0° F., or about – 18° C. The freezing point of sap in plants or trees is lower than that of pure water. This is one reason why the temperature must fall several degrees below 32° F., or 0° C., before fruit buds and other vegetation are injured by freezing.

For dilute solutions the depression of the freezing point is proportional to the percentage of the dissolved substance.

The low temperatures which can be attained by use of *freezing mixtures* are due to the fact that the freezing points of solutions are lower than that of pure water. When a liquid and its solid (for example, pure water and ice) are together, the temperature always falls to the freezing point. Similarly, when ice is placed in a solution of water and common salt, some of the ice melts, and the temperature falls to the freezing point of the solution. With a sufficient amount of salt and well-crushed ice a temperature of about 0° F. can be obtained. If four parts of crystalline calcium chloride are mixed with three parts of snow, a temperature of about – 50° C. can be produced.

266. Evaporation; the kinetic theory. The other change of state is the change from a liquid to a vapor or from a vapor to a liquid. We all know that water can change into vapor; and we have evidences in rain, dew, and the moisture on the outside of cold vessels that water vapor can change back into a liquid. As has been stated before, it is important for the student to form a rather clear mental picture of the kinetic theory of gases and vapors. In the case of water vapor we must imagine enormous numbers of high-speed water molecules darting in every direction. These particles are so small that, notwithstanding their enormous numbers, they are, on the average, relatively far apart. Some idea of this condition is obtained when one learns that 1 cubic centimeter of water forms 1670 cubic centimeters when changed into vapor (steam) at 100° C. and a pressure of 76 centimeters. The student should recall also that the average speed of molecules depends on the temperature: the higher the temperature, the greater the average speed.

Evaporation takes place because the high-speed molecules have a tendency to escape through the water surface. When water is first put into a closed vessel, partly filling it, the density of the vapor in the upper part of the vessel increases rapidly, owing to the escape of the molecules from the water. But another process is going on: molecules are continually going from the vapor back into the liquid. As the density of the vapor increases, the rate at which molecules go back into the water increases. The density will increase until the number of the molecules returning to the water each second is the same as the number leaving it each second. When the density of the vapor has reached this value, there is said to be a condition of *equilibrium*. When water and its vapor are together, there is thus a continual exchange going on: molecules are escaping from the liquid, and molecules in the vapor are returning into the liquid. It is important that the student understand clearly this idea of an equilibrium condition, for it will materially aid in understanding what follows.

Water changes into vapor at all temperatures. Even ice evaporates. It is well known that wet clothing hung out in freezing weather will dry after it freezes. In many parts of America it is not an unusual thing for snow to disappear under the influence of a dry wind even when the temperature is much below freezing. On the other hand, water vapor will pass over into a solid state without going through the liquid stage. The formation of frost and snow offers instances of the change of vapor directly into a solid form. The process in which a solid changes into a vapor or a vapor into a solid is called *sublimation*. Camphor, iodine, and other substances will change directly from a solid into a vapor, and vice versa.

267. Saturated vapors. When a vapor in contact with its liquid has reached that maximum density where the vapor is in equilibrium with the liquid, it is said to be a *saturated vapor*. Stated in terms of our molecular picture, the density of a saturated vapor is such that the number of molecules returning each second to the liquid is equal to the number leaving the liquid.

The pressure of any saturated vapor is a constant, provided the temperature does not change. The pressure cannot be altered by changes in the volume of the vapor. It can be proved by direct experiment that this theorem is true. But the following theoretical proof should be carefully studied:

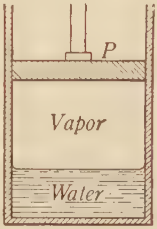


FIG. 165

Consider a liquid and its vapor in contact, as shown in Fig. 165. The top of the vessel is a piston which can be moved up and down, thus changing the volume occupied by the vapor. The pressure produced by any vapor or gas is caused by a bombardment of the walls by the rapidly moving molecules. The magnitude of the pressure depends on the number of molecules and on their speed. The first of these factors (the number of molecules) depends on the *density* of the gas or vapor. The average speed depends on the temperature. If the temperature is kept constant, the pressure of a gas or vapor can change only as its density changes. If it can be shown that the density of a saturated vapor cannot be increased, except momentarily, then we have proved that the pressure cannot be increased. If the piston of Fig. 165 is moved down, the density of the vapor may be momentarily increased. If so, equilibrium will be destroyed; for the rate at which molecules go from the vapor into the liquid is increased and will be greater than the rate at which molecules leave the surface. Thus the density will decrease quickly until it is again the same as it was before the piston was pushed down. Hence pushing the piston down does not increase the density or the pressure of the saturated vapor.

If the piston of Fig. 165 is raised, the density and pressure of the vapor will momentarily be lowered; but the evaporation of the liquid will quickly raise the density and pressure back to the equilibrium condition of the saturated vapor. Hence *the pressure of a saturated vapor cannot be changed by changes in the volume when the temperature does not change.*

A simple experiment brings out the differences in the behavior of a saturated vapor and a gas. Two glass tubes, each closed at one end, are filled with mercury and inverted in a mercury "well" (Fig. 166). By means of a small eye dropper or fountain-pen filler a little ether is injected into the mercury below the end of one of these tubes. The ether will bubble up through the mercury into the tube and float on top of the mercury. (Considerable care must be taken to prevent air from going in with the ether.) The space above the mercury will quickly fill with saturated ether vapor. Since the pressure of saturated ether vapor at room temperatures is considerable, the mercury surface in this column will at once fall an appreciable distance. Some air should be allowed to bubble up into the other tube. The experiment consists in watching the changes in the pressure and volume of the air and of the ether vapor when the tubes are pushed down into the well or raised up. The pressure inside one of these tubes is equal to the height of the mercury column in a barometer *minus* the height of the column of mercury in the tube. Hence changes in pressure inside these tubes will be shown by changes in the heights of the mercury columns. When the tube containing the ether is slowly raised, the surface of the mercury in the tube remains nearly at the same height, although the volume occupied by the vapor may be increased a great deal; when the tube is lowered, the vapor condenses as its volume decreases, but with scarcely any change in the height of the mercury column. If the tube is free from air, and the changes take place so slowly that there are no temperature changes, the pressure of the vapor will not change. But evaporation and condensation, which take place when the volume changes, are cooling and heating processes respectively; thus there probably will be slight changes in the temperature, and hence small changes in the height of the mercury column. These changes will be insignificant, however, compared with the changes that take place in the tube containing air. In this case the pressure varies inversely as the volume, obeying Boyle's law. When the volume is doubled, the pressure is half as great. This experiment and many others show that when the temperature remains constant, *the pressure of a saturated vapor is independent of its volume*

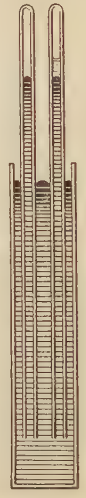


Fig. 166

268. Nonsaturated vapors. In the experiment of Fig. 166, if there had been only a small amount of ether in the tube, and it had evaporated before the tube was fully raised, the vapor would not have remained saturated and its pressure would have changed. As a vapor gets away from a saturated condition it tends to obey Boyle's law. At a considerable distance from a saturated condition it has all the properties of a gas, and may be called a gas. All gases can be saturated and liquefied at sufficiently low temperatures. When near saturation they are called vapors. For example, air at a pressure of 76 centimeters and a temperature of about − 190° C. is a *saturated* vapor, no longer obeying Boyle's law and other gas laws.

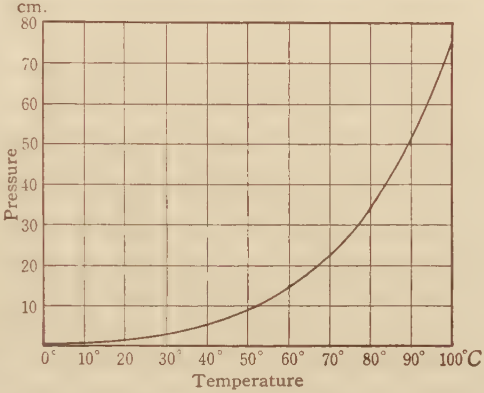


Fig. 167. Pressure of saturated water vapor

269. Dependence of pressure of saturated vapor on temperature. Imagine a closed airtight vessel containing some water, with no air (only water vapor) in the space above the water. If the temperature of this vessel and its contents is raised, the pressure of the saturated vapor is increased. This can be determined by direct experiment. The observed pressure of saturated water vapor at various temperatures is shown in the graph of Fig. 167 and also in the tables on pages 280 and 296.

In general, the pressure of any vapor in contact with its liquid will change as the temperature changes. It is helpful to interpret this fact in terms of the molecular theory. When the temperature is raised, the tendency of molecules to leave the surface of the liquid is increased. This tends to destroy the equilibrium between the number of molecules leaving the surface and the number of those returning to it. The density of the vapor will then increase. But as the density increases, the number of molecules returning to the surface increases; hence equilibrium will soon be restored. But since an increased density means a larger vapor pressure, equilibrium is restored at an increased vapor pressure.

270. Boiling. When a liquid is heated to a certain temperature, bubbles filled with saturated vapor will form in the interior of the liquid and rise to the surface. In this case the evaporation of the liquid is taking place not only from its top surface but also in the interior. Such a process is called *boiling*. The student should keep in mind the fact that these bubbles are filled with saturated vapor, and that the pressure inside them is determined by the temperature of the liquid. When water is first heated, small bubbles of the air which has been in solution form on the walls of the vessel. It is very easy, however, to distinguish between small air bubbles and those filled with saturated vapor.

That boiling takes place at a definite temperature is one of the noteworthy facts of nature. So definite is the temperature of pure, boiling water at normal pressure that it has been taken universally as one of the standards of temperature. It makes little difference whether the water is boiling slowly or rapidly. A better knowledge of this fact should greatly reduce gas bills where cooking is done by gas. The fact is that when the liquid begins to boil, the maximum temperature has been reached, and nothing is gained by rapid boiling. When the liquid begins to boil, the gas supply should be reduced, and merely enough heat furnished to keep the liquid at the boiling temperature.

271. Effect of pressure on the boiling point. When water to which heat is being supplied is confined in a closed vessel so that the steam, or water vapor, cannot escape, the pressure inside increases and the temperature of the boiling water becomes higher. For example, in a steam boiler, when the pressure rises to 40 pounds per square inch (25 pounds per square inch above atmospheric pressure), the water is boiling at a temperature of about 287° F. instead of 212° F., the boiling temperature for atmospheric pressure. *Increasing the pressure raises the boiling point of water.*

Cooking takes place more quickly and thoroughly at higher temperatures. In "pressure cookers" the higher temperature is secured by boiling water under pressure.

A flask (see Fig. 168) is partly filled with hot water and is tightly fitted with a rubber stopper. A glass tube which runs through the stopper into the space above the water is connected to a filter pump, or aspirator. When the pressure of the air above the water is sufficiently reduced by the pump, the water will begin to boil. With a good pump the boiling will continue until the temperature of the water falls nearly to room temperature. *Decreasing the pressure lowers the boiling point of water.*

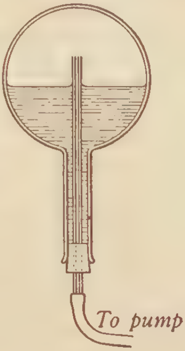


FIG. 168

The ordinary fluctuations of atmospheric pressure at any one place cause changes of about 2° C. in the boiling point of water, which is 100° C. only when the barometer is 76 centimeters high,—a height seldom reached at places much above sea level. The following table shows the variation in the average pressure and boiling point of water with elevation:

ELEVATION IN FEET	MEAN BAROMETER		MEAN BOILING POINT	
	Inches	*Centimeters*	*F.*	*C.*
0	29.9	76.0	212°	100°
1000	28.8	73.2	210°	99°
5000 (Denver)	25.0	63.5	203°	95°
10,000	20.8	52.8	194°	90°
14,000 (Pikes Peak)	17.9	45.6	187°	86°

272. Why the boiling point depends on pressure. By the aid of facts which he has learned about the pressure of saturated vapor, the student may already have explained to his own satisfaction why the temperature at which a liquid boils depends on the pressure exerted on the surface. If not, he can easily do so by recalling three facts:

1. The bubbles in a boiling liquid are filled with saturated vapor.
2. The pressure of the vapor inside the bubbles must be slightly greater than the pressure on the top surface of the liquid. However, at small depths, in the interior of the liquid the pressure is very little different from that on the surface.

3. The pressure of a saturated vapor of any liquid changes only when the temperature changes.

To apply these facts let us consider some water in a closed vessel where the pressure, by some suitable means, is kept constant at 148.9 centimeters of mercury (about 2 atmospheres). Let us gradually raise the temperature of the water. At low temperatures it is impossible for vapor bubbles to form in the interior, for the pressure of saturated water vapor at low temperatures is so small that the bubbles would be crushed by the pressure which is transmitted throughout the liquid. An inspection of the table given in this section shows that the pressure of saturated water vapor is 148.9 centimeters when the temperature is 120° C. Hence when the vessel and its contents reach 120° C., the pressure of saturated vapor will be equal to the pressure on the surface of the water, and boiling can commence.

Boiling takes place at that temperature at which the pressure of saturated vapor is equal to the pressure on the surface of the liquid.

If the pressure on the water had been only 35.5 centimeters, then the water would have boiled at a temperature of 80° C., for at that temperature the pressure of saturated vapor is equal to 35.5 centimeters of mercury (see the table given below). It should now be obvious that since water boils at 100° C. under a pressure of 76 centimeters of mercury, the pressure of saturated water vapor at 100° C. must be 76 centimeters.

PRESSURE OF SATURATED WATER VAPOR AT DIFFERENT TEMPERATURES; ALSO PRESSURE UNDER WHICH WATER WILL BOIL AT DIFFERENT TEMPERATURES*

Temperature in Degrees Centigrade	Pressure in Centimeters of Mercury	Temperature in Degrees Centigrade	Pressure in Centimeters of Mercury
0	0.46	60	14.9
10	0.92	80	35.5
20	1.75	100	76.0
30	3.17	120	148.9
40	5.51	140	270.9

*See also the graph of Fig. 167 and the table on page 296.

The foregoing table might be called a *steam* table, for it gives the pressure of the steam (or saturated vapor) in a boiler for different temperatures of the water.

273. Pressure of saturated vapors of different liquids. The pressure of saturated vapor of different liquids varies through a wide range. In many physical laboratories there is kept a group of barometers, standing in a trough of mercury, to show the difference in the pressures of different saturated vapors at room temperatures. The first tube usually has a vacuum* in the top; hence the height of this column indicates the true barometric pressure. In the second there may be a little water in the top. If so, this column stands about 1.7 centimeters lower than the first, showing that at room temperature (assumed to be 20° C.) the pressure of saturated water vapor is about 1.7 centimeters. In the top of the third there will be a different liquid, say alcohol, and in the fourth perhaps a little ether. The heights of the mercury columns show that the vapor pressure of alcohol, at room temperature, is over twice that of water, and that of ether considerably more.

274. The boiling points of different liquids. Since the pressures of the saturated vapors of different liquids at the same temperature are different, it follows that the *temperatures* at which the pressures of the saturated vapors will be equal to *atmospheric pressure* are different for different liquids. Hence the temperatures of their boiling points are different.

TABLE OF BOILING POINTS AT A PRESSURE OF 76 CENTIMETERS OF MERCURY

	Degrees Centigrade		Degrees Centigrade
Alcohol (ethyl)	78.3	Naphthalene	218.
Ammonia	−33.5	Nitrogen	−196.
Ether	34.6	Oxygen	−183.
Hydrogen	−253.	Sulfur	444.6
Lead	1525.	Tin	2270.
Mercury	356.7	Turpentine	159.

---

*In reality it contains saturated mercury vapor at a pressure of about 0.0001 centimeter of mercury. This is so small that it is negligible.

275. The pressure exerted by a vapor is independent of the presence of other gases or vapors. A law discovered by Dalton states that *the pressure exerted by a vapor is the same whether other gases or vapors are present or not*, provided there is no chemical action between the different gases and vapors.

Ordinary air is a mixture of gases and vapors, with water vapor usually an important constituent. Dalton's law tells us that the presence of these other gases does not change the pressure exerted by the water vapor. What we call atmospheric pressure is merely the sum of the pressures which the gases and vapors making up the atmosphere exert separately. If the water vapor is saturated on a damp day when the temperature is 20° C., its pressure, according to the table on page 280, is 1.75 centimeters. The rest of the atmospheric pressure is due to the other constituents —nitrogen, oxygen, carbon dioxide, etc. It is incorrect to think of the air as a sort of sponge which can take up so much vapor. The density and pressure of a saturated vapor depend only on temperature and have nothing to do with the presence of other gases.

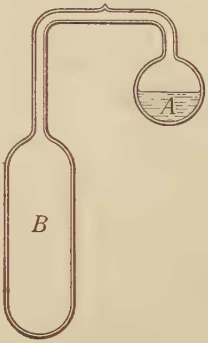


FIG. 169

276. Distillation. The fundamental principle involved in *distillation* may be illustrated by the aid of the simple piece of apparatus shown in Fig. 169. A wide glass tube has two bulbs, *A* and *B*, on its ends. If a little water is placed in *A* and the tube sealed up, the space in *A* and *B* and the connecting tube will soon be filled with saturated water vapor. If the temperature of the entire vessel and contents is 20° C., then the pressure of this saturated vapor is about 1.7 centimeters. But let *B* be kept at a temperature of 10° C., while *A* is kept at 20° C. The pressure of water vapor in *B* can no longer be 1.7 centimeters; it cannot be larger than that corresponding to the pressure of a saturated vapor at 10° C., which, according to the table on page 280, is 0.92 centimeter. Hence the vapor pressure in *B* will be less than that in *A*. Vapor will then move from the region of higher pressure over to *B*, where it will condense. The result is that the water in *A* will evaporate and the vapor, passing over into B, will condense. Such a process as this is called *distillation.*

The formation of moisture on cold vessels or of dew on cold objects near the ground is a condensation resulting from the cooling of these things below the temperature corresponding to the pressure of the vapor.

An interesting case of diffusion and condensation, or distillation, of water vapor takes place in soils. The tendency is for the water vapor in the small cavities of the soil to be saturated. But if the soil is not of uniform temperature, a process something like that demonstrated in connection with Fig. 169 takes place. For example, if on a clear, cold night the top layer of the soil becomes cold, water vapor from the warmer soil below diffuses up and condenses in the colder part. The increase in the moisture of the surface layer is often quite noticeable in the morning. In the daytime, when the surface layer is warmer than the deeper part, the water vapor tends to diffuse down and condense in the cooler layers below. In the early spring, while the temperature a short distance below the surface is low, the soil does not dry out so quickly after a rain as it may in the fall, when the lower layers are warmer. Why?

277. Fractional distillation. If a mixture of liquids is heated, vapors of all the liquids will be given off, but not in equal amounts. The more volatile liquid, the one with the lower boiling point, will give off vapor more rapidly than the others. For example, if a mixture of 50 per cent of water and 50 per cent of alcohol is heated, both water and alcohol vapors are given off, but more alcohol than water. If these vapors were drawn into a cold chamber, where they would condense, more alcohol than water would be in the distillate. If this distillate were again heated and the vapors carried off and condensed, the percentage of water in the second distillate would be found to be still less. This process is called *fractional distillation.*

Crude petroleum is a mixture of many substances having different boiling points. The process of refining is largely one of utilizing the method of fractional distillation. When the crude oil is heated, the more volatile substances come off more rapidly. By means of fractional distillation the mixture is separated into many different grades of oils and solids.

278. The boiling points of solutions. The boiling point of a solution (for example, water and sugar) is higher than that of the pure solvent. The amount the boiling point is raised depends on the concentration; that is, on the relative amount of the substance dissolved.

The boiling of sirup diminishes the amount of water in it, for the escaping vapor carries off water but no sugar. Hence, as the boiling continues, the concentration increases, and the temperature at which the sirup boils rises. In the making of candies the sirup is boiled until the desired concentration is reached, different kinds requiring different concentrations. A thermometer, by which the temperature of the boiling liquid is measured, may be used to tell when the desired concentration is reached. For example, for making one kind of candy, we may be told to boil the sirup until the temperature reaches 110° C.; for making a different kind a temperature of 120° C. must be reached.

Since the boiling point of a solution is higher than that of pure water, it follows that the pressure of the saturated vapor given off by a solution is less than the vapor pressure of pure water at the same temperature.

279. Heat of vaporization. We have seen that energy in the form of heat must be added to ice to change it into water. Similarly, it is found that heat energy must be added when water is changed into water vapor.

The heat of vaporization is equal to the number of calories of heat required to change 1 gram of liquid into vapor without any temperature change.

It is found by experiment that the heat of vaporization of water at 100° C. is 539 calories per gram.

When water evaporates at a temperature lower than 100° C., the heat of vaporization is somewhat higher. At 0° C. it is about 600 calories per gram.

The fact that the heat of vaporization of water is a large quantity becomes evident if one estimates what 539 calories will do. For example, the heat required to change 1 gram of water at 100° C. into vapor is more than sufficient to raise 5 grams of water from 0° C. to the boiling point, and enough to melt 3 grams of ice at 0° C. and raise it to 100° C.

When water vapor condenses, heat is liberated in an amount equal to that required to evaporate an equal mass of water.

280. Evaporation a cooling process. Whenever the heat required to evaporate water is supplied by the water or by adjacent bodies, evaporation is a *cooling process*. Since the heat of vaporization of water is so large, the cooling effect when water vaporizes is large. A cloth or a sheet of filter paper when wet with water and shaken out will quickly become colder than the air, except when the water vapor in the air is so nearly saturated that drying takes place very slowly. In the central part of the United States it is not an uncommon thing in summer for wet clothing in the wind to fall 20° F. below air temperature.

The evaporation of the perspiration on our bodies is a cooling process. All who live where it gets hot in summer know the difference in comfort between a dry day and one when the humidity is so high that evaporation does not take place rapidly. It is not an uncommon thing for people to live where the temperature for many hours at a time is higher than body temperature. The evaporation from their bodies keeps them cooler than the air. In dry climates the temperature of bottled milk will remain lower than air temperature if the bottles are surrounded by damp cloths. The so-called *iceless refrigerator* is a framework covered with cloths which are kept wet. This is efficient when the air is relatively dry and in motion. In tropical countries water is often kept in unglazed earthenware vessels in a place where wind can blow on them. The water, slowly seeping through the pores, evaporates from the outside of the vessel, with the result that the water inside is kept cool. Automobile tourists often carry drinking-water in canvas bags, which permit a slow seepage of water to the outside. It is observed that the temperature of a wet soil does not rise as rapidly in the spring as in a well-drained soil. This is due partly to the increase in thermal capacity produced by the presence of water, but to a greater extent to the cooling which is caused by evaporation.

281. Freezing water by its evaporation. The following is a very interesting and striking experiment: An air pump with a bell jar having a relatively small volume is used. On the plate of the pump is placed a glass dish containing some strong sulfuric acid. Over this dish, in a much smaller receptacle, is placed a little cold water. It has been explained before that water will boil whenever the pressure on its surface is reduced to a pressure equal to that of the saturated water vapor (see tables in sections 272 and 291). When the pump is operated, and a sufficiently low pressure is reached, the water will begin boiling. The acid absorbs water vapor very rapidly and helps the pump to maintain a low pressure under the bell jar. The boiling of the water means a very rapid evaporation, with a large cooling effect. Hence the water gets colder, but it will continue to boil if the pump keeps reducing the pressure. The water thus gets colder and colder until its freezing point is reached, when it will begin to freeze. *It freezes and boils at the same time.* The lowering of the temperature and the freezing are caused by the removal of the heat which is required to change the liquid into vapor. The heat of vaporization is so much greater than that of fusion that much more water will be frozen than evaporated; for example, if, while the temperature is at the freezing point, 1 gram of water evaporates, it will take away 600 calories,—enough to freeze 600/80, or 7.5, grams of ice.

The temperature at which a liquid can both boil and freeze at the same time is called the *triple point.* All three states of the substance are in equilibrium at that temperature. The triple point of water occurs at a pressure of 0.46 centimeter of mercury and a temperature of $+0.0072°$ C. (the temperature is not 0°, because of the decreased pressure).

282. Carbon-dioxide snow. The following experiment gives an excellent illustration of the cooling effect produced by rapid evaporation combined with an expansion of vapor. Carbon dioxide in liquid form under high pressure is sold commercially in steel tanks, or drums. Generally the drum is used in an upright position, and the vapor is drawn off from the top through a valve. If the drum is inverted, and the valve opened, the liquid at the bottom will be forced out. But carbon dioxide will stay in a liquid form at ordinary temperatures only under enormous pressure; hence the liquid immediately evaporates, and the vapor expands, both being cooling processes. The result is that such a low temperature is reached that some of the vapor will be frozen in the form of a white snow. If the valve is surrounded by a cloth bag, the snow is easily caught. This snow, temperature $-78°$ C., will not melt, but will slowly evaporate, retaining its low temperature by the cooling effect of its evaporation. If a little of this snow is mixed with ether, a very cold freezing mixture is produced. By its aid mercury can be frozen and other interesting experiments performed.

283. Critical temperatures. A simple experiment is of great help in understanding some important facts about what is called *critical temperature*. The glass tube illustrated in Fig. 170 is about half full of liquid carbon dioxide, the rest of the tube being filled with the saturated vapor. At a room temperature of 20° C. the pressure of the saturated vapor of carbon dioxide is about 60 atmospheres, or about 900 pounds to the square inch. Hence the pressure in the tube is enormous; and suitable precautions should be taken, because these tubes sometimes explode. If the temperature of the tube and its contents is slowly raised, the liquid expands rapidly and its *density* becomes *less*. On the other hand, the pressure and density of the saturated *vapor* become *greater*. Hence, as the temperature rises the density of the vapor and that of the liquid tend to become equal. At a temperature of about 31° C. the two densities become equal, and there is no longer any difference between the vapor and the liquid. By watching the tube as the temperature rises it will be observed that the surface of separation (the meniscus) between the liquid and the vapor gradually *disappears*. This temperature, at which the two densities become equal, is known as the *critical temperature*.

FIG. 170

At a temperature higher than its critical temperature no substance can exist as a liquid, no matter what the pressure may be.

CRITICAL TEMPERATURES AND PRESSURES

	DEGREES CENTIGRADE	ATMOSPHERES
Air	− 140°	39.
Ammonia	132°	115.
Carbon dioxide	31.3°	73.
Ether	194°	35.6
Hydrogen	− 241°	20.
Nitrogen	− 146°	33.
Oxygen	− 118°	50.
Water	365°	200.

The critical temperature of water is 365° C. Water vapor cannot be liquefied at a temperature higher than this. The critical temperature of air is $-140°$. Unless it be cooled below that temperature it cannot be liquefied, no matter how much pressure is applied.

Critical pressure is the pressure of the saturated vapor at the critical temperature.

284. The spheroidal state. When water is dropped on a very hot plate, such as the top of a stove, it collects in drops which dance around vigorously. Careful observation shows that the drops do not touch the hot metal, but are supported by a layer of vapor. If the student places his eye in the same plane as the hot surface, he will be able to see between the drop and the surface. When a liquid is supported by a layer of its vapor, it is said to be in the *spheroidal state.*

When one touches a very hot metal with a moistened finger, it is the layer of vapor from the rapidly evaporating water that prevents the finger from being burned. A wet hand can be plunged for an instant into molten metal without injury. Pouring liquid air at its very low temperature into any receptacle, or even on the hand, is similar to pouring water into a red-hot vessel. The liquid air does not touch the hand or the vessel, being separated from it by a layer of vapor (air).

285. Ice and refrigerating machines. Advantage is taken of the cooling produced by the rapid evaporation and expansion of ammonia in the operation of commercial refrigerating and ice machines. The principle involved can be readily understood by the aid of Fig. 171. This diagram does not give the actual details and arrangement of the apparatus, but only those features necessary for an understanding of the principles. A pump draws ammonia vapor from the side *A* and forces it under pressure into the pipes which run through the tank *B*. The compression heats the gas, but it is cooled in the tank *B* by water which circulates over the pipes. Under the effects of the pressure, and the lowering of temperature by the water, the vapor in these pipes liquefies. For example, at a temperature of 20° C. ammonia is liquefied by a pressure of about 8.5 atmospheres. The liquid ammonia escapes from these pipes through a valve *V* into the pipes on the left. The pressure in these pipes is kept low, for the compressor *P* is continually pumping the ammonia out of this side and forcing it toward the right. When the liquid ammonia escapes into this region of low pressure, it rapidly evaporates and expands. Since these are cooling processes, the expanded vapor and the pipes on the left get very cold. These pipes cool the brine (*A* in Fig. 171) which circulates through the refrigerator, thus keeping it cold; or, if used in the manufacture of ice, the pipes may pass through a large tank of brine which is cooled down to a low temperature. Ice is formed when metal containers filled with water are placed in the cold brine.

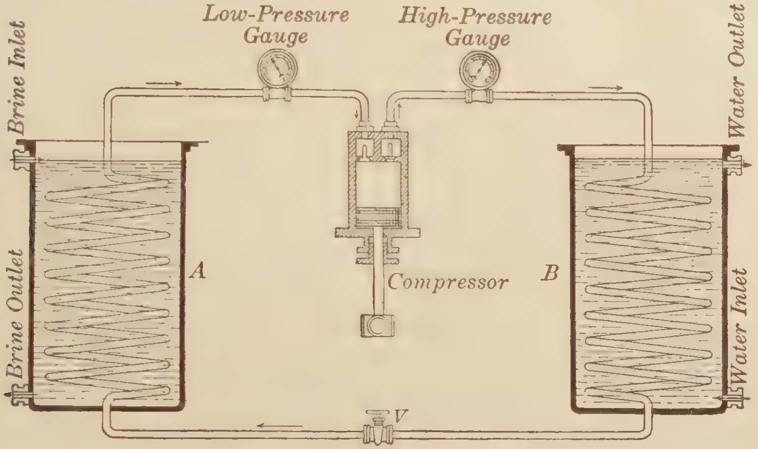


FIG. 171. Refrigerating machine

In this process heat is extracted from one region (tank *A*) and given up to another (tank *B*). But although heat energy is taken away from the refrigerator or the brine, this energy cannot be used to run the compressor. According to the second law of thermodynamics (sect. 234) heat energy will not of itself go from a cold body to a warmer one and operate any machine: energy from *outside* must be used to take heat out of a cold body.

286. Liquefaction of air. The process of liquefying gases is now on a commercial basis. For example, oxygen in large quantities is obtained by liquefying air and then allowing the nitrogen, which boils at a lower temperature, to boil off.

The principle commonly employed is easily explained. Air is compressed to 150 or 200 atmospheres. Since the compressing of this air heats it, it is usually cooled to room temperature by water which circulates through the compressor. Sometimes a freezing mixture of salt and ice is used to cool the air further. The cooled compressed air then passes into a coil of pipe on the inside of the liquefier (Fig. 172). The liquefier must be well insulated from the heat of the room. Evacuated double-walled Dewar flasks, as represented in the figure, are often used. At the end of the coil of pipe in the liquefier is a fine valve through which the compressed air escapes. The rapidly expanding air is cooled by expansion (the Joule-Thomson effect), and then flows up over the coils of pipe, out at *A*, and back again to the compressor. Since this cold air flows over the pipes in the liquefier, it cools the compressed air inside them, so that when this air in turn expands, a lower temperature is reached. This colder air in turn, rising up through the liquefier, lowers the temperature of the pipes and of the inclosed compressed air still further, and so on. In this way the expanding air gets colder and colder, until some of it liquefies. The liquefied air falls to the bottom and can be drawn off into a suitable receptacle.

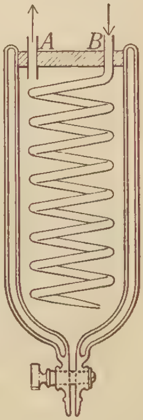


FIG. 172

287. Summary. The melting and solidifying of many substances take place at a definite temperature, but in some cases (butter, for example) the change in state is a gradual one.

When a substance changes from a solid to a liquid, the internal energy is increased, and heat must be supplied to bring about the change. If heat is added to a mixture of ice and water, ice will be melted; but if heat is taken away, water freezes.

The *heat of fusion* of any substance is equal to the number of calories of heat required to melt 1 gram of the substance, its temperature remaining unchanged. When a gram of a substance solidifies, it gives up calories equal to the heat of fusion. The heat of fusion of ice is 79.8 calories per gram of ice melted.

Some substances expand on solidifying, others contract. An increase of pressure lowers the temperature of the melting points of those substances that expand on solidifying and raises the melting points of those that contract on solidifying. The phenomenon of *regelation* is due to the change of melting point through pressure.

Sometimes a liquid (pure water in a clean vessel, for example) may be cooled below its freezing point without freezing. This is called *undercooling*.

When a substance is dissolved in water, the freezing point of the solution is lower than that of pure water.

When a vapor in contact with its liquid has reached its maximum density, it is said to be *saturated*. The pressure of a saturated vapor can be changed only by a change in its temperature.

The *boiling* of any liquid takes place at a definite temperature for a definite pressure. An increase of pressure raises the boiling point, and a decrease lowers it. Boiling takes place at that temperature at which the pressure of the saturated vapor is equal to the pressure on the surface of the liquid.

Since the pressures of the saturated vapors of different liquids at the same temperature are different, it follows that the temperatures at which the pressures of saturated vapors will be equal to atmospheric pressure are different for different liquids. Hence the temperatures of their boiling points are different.

The *heat of vaporization* is equal to the number of calories of heat required to change 1 gram of liquid into vapor without any temperature change. The heat of vaporization of water at 100° C. is 539 calories per gram.

Evaporation is a cooling process; condensation, a heating process.

Water may be frozen by its evaporation. At the *triple point* a liquid may freeze and boil at the same time.

At a temperature higher than its *critical* temperature no substance can exist in liquid form, no matter what the pressure may be.

The cooling effect of the evaporation and expansion of ammonia is used in refrigeration. In a liquid-air machine the cooling due to the free expansion of compressed air is utilized.

PROBLEMS

1. 25 gm. of ice at 0° C. are added to 400 gm. of water at 40° C. Compute the resulting temperature.

2. 200 gm. of water at a temperature of 30° C. are poured into a cavity in a block of ice at 0° C. How much ice should be melted?

3. A bullet of mass 9 gm. and having a speed of 700 m./sec. is shot into a block of ice. If the bullet before entering was at the temperature of the ice, 0° C., how much ice should be melted by the energy of the bullet?

4. A minute crystal of ice is dropped into 160 gm. of water which has been undercooled to − 5° C. The temperature immediately rises to 0° C. How much ice will be formed?

5. A liter of water (1000 gm.) is cooled from 80° C. to a final temperature of 20° C. by adding snow and by stirring. How much snow is required?

6. In a determination of the specific heat of iron a mass of 160 gm. was heated to 100° C. and dropped into a cavity in a block of ice. The mass of ice melted was 22.4 gm. Compute the specific heat of iron.

7. 250 gm. of ice at 0° C. are dropped into a copper calorimeter the thermal capacity of which is 18 and which contains 400 gm. of water at 60° C. Find the resulting temperature.

8. A mass of 400 gm. of copper (specific heat 0.093) is heated in an oil bath and then placed in a cavity in a block of ice. The mass of ice melted is 75 gm. Required to find the temperature of the bath.

9. 1 kg. of lead (specific heat 0.031) is dropped into a vessel (thermal capacity 20) which contains 200 gm. of water and 50 gm. of ice at 0° C. The initial temperature of the lead is 300° C. Find the temperature of the mixture.

10. How many calories are required to change 50 gm. of water at 50° C. to steam at 100° C.?

11. 30 gm. of steam at 100° C. are passed into a cavity in a large block of ice at 0° C. How many grams of ice will be melted?

12. A calorimeter (specific heat 0.093) which has a mass of 90 gm. contains 391 gm. of water at 5° C. 20 gm. of steam are forced in, and the temperature becomes 35° C. Compute the heat of vaporization of the water.

13. 10 gm. of steam at 100° C. and 50 gm. of ice at 0° C. are added to 400 gm. of water at 30° C., contained in a 200-gram copper calorimeter. Compute the final temperature of the mixture. (The specific heat of copper is 0.09.)

14. 40 gm. of ice at 0° C. and 10 gm. of steam at 100° C. are added to a 100-gram copper calorimeter containing 300 gm. of water at 20° C. Find the resulting temperature. (The specific heat of copper is 0.09.)

15. Compute the number of heat units required to raise 20 gm. of ice at − 15° C. to steam at 150° C. (The specific heat of both steam and ice is 0.5.)

16. What amount of steam at 100° C. must be passed into 16 kg. of water at 0° C., in which 4 kg. of ice are floating, in order to raise the whole to 30° C.?

CHAPTER XXII

SOME APPLICATIONS TO METEOROLOGY

Dew and frost, 288. Experimental methods of finding the dew point, 289. Humidity, 290. Methods of measuring humidity, 291. Fog and cloud, 292. Rain or snow, 293. Conditions causing clouds or rain, 294. Winds: general causes, 295. Effect of the rotation of the earth, 296. Cyclones, 297. Anticyclones, 298. Effect of cyclones and anticyclones on the weather, 299. Thunderstorms, 300. Tornadoes, 301. The averaging of temperature, 302. The averaging of rainfall, 303.

The science of meteorology deals with those phenomena of the atmosphere which affect our weather. To a very large extent it is a series of applications of the principles of physics to these special problems. So true is this that meteorology has been defined as the physics of the atmosphere. Obviously in a text of this kind it would be out of place to attempt to cover much of this important and extensive subject. However, some of the more commonly observed phenomena, which are excellent illustrations of principles explained in this book, will be given in this chapter.

288. Dew and frost. When cold water is poured into a glass on a warm, humid day, the vapor in the air near the glass will condense on the glass, provided the temperature of the cold water is low enough to cool the air next to the glass below the point at which the water vapor in the air becomes saturated. At night, objects near the ground frequently cool by radiation until they are colder than the air, and water vapor of the air condenses on them in the form of dew or frost.

The pressure of saturated water vapor at 0° C. is 0.46 centimeter of mercury. When the air contains water vapor the pressure of which is less than 0.46 centimeter, the vapor must be cooled to some temperature *below* the freezing point before it will be saturated. If it came in contact with some body of sufficiently low temperature so that some of the vapor condensed on the body, it would not form drops of water but would form *frost*.

The *dew point* of the atmosphere is defined as that temperature to which it must be lowered in order that the water vapor which it contains should become saturated. If the pressure of the water vapor in the air is 0.92 centimeter, then, no matter what the temperature of the air may be, its dew point is 10° C., or 50° F. For at this temperature (see the table on page 296) water vapor at a pressure of 0.92 centimeter is saturated.* When the dew point is below freezing, it is often called the *frost point*.

Dew or frost collects on objects if their temperature goes below the dew point of the air. In cold weather water or frost often collects on the inside of our windows. This is because the temperature of the glass is below the dew point of the air in the house. If dew or frost does not collect on the windows in severe winter weather, it usually indicates that the air in the house is very dry, with a low dew point.

289. Experimental methods of finding the dew point. The following is a very simple method for finding the dew point: A thin-walled polished metallic cup, a thermometer, and some ice water are needed. Water is poured into the cup, and its temperature is gradually lowered by adding small quantities of ice water while stirring thoroughly. When a film of moisture begins to appear on the cup, the thermometer is read. Then the temperature of the cup is allowed to rise slowly. When the film begins to evaporate, the temperature is read again. The mean of these two readings gives the dew point. The more slowly the temperature is lowered in the neighborhood of the dew point, the more accurately can the dew point be determined. Practically all the direct methods are modifications of this simple one. The dew point is usually obtained in meteorological stations from the readings of the wet-bulb and dry-bulb thermometers. This purely empirical method will be explained later.

---

*The change in pressure of the vapor due to the change of temperature is neglected. The pressure will not stay exactly at 0.92 centimeter when the air is cooled, but this change is relatively unimportant unless the change in temperature is very large.

290. Humidity. The *absolute humidity* of the atmosphere is defined as the mass of water vapor present in a unit volume. Usually the humidity is expressed in grams per cubic meter or in grains per cubic foot. The amount of water vapor present in the air varies a great deal, depending on the temperature and degree of saturation. When air is at a temperature of 0° C., and the water vapor is saturated, there are about 4.8 grams per cubic meter; but if the temperature is 30° C. (86° F.), and the amount of water vapor is only half of that required to produce saturation, there will be about 15 grams per cubic meter (see the table on page 296).

We are usually much more interested in what we sometimes call the dampness of the air. This is a relative term, depending on the degree of saturation. For example, if the air temperature is 0° C. and there are 4.7 grams per cubic meter of water vapor present, it is a very damp day; but if the temperature were 30° C. (86° F.) and there were only 4.7 grams per cubic meter of water vapor present, it would be an extremely dry day, for at that temperature it requires about 30 grams of water vapor per cubic meter to produce saturation (see the table on page 296). The scientific term for stating the relative dampness of the air is *relative humidity. The relative humidity is defined as the ratio of the absolute humidity to that required to produce saturation.*

Since the mass of water vapor present in the air is closely proportional to the pressure of the vapor, it follows, approximately, that, for any definite temperature,

$$\frac{\text{Mass of vapor per cu. m.}}{\text{Mass of saturated vapor per cu. m.}} = \frac{\text{pressure of vapor}}{\text{pressure of saturated vapor}}.$$

Hence the relative humidity may also be defined as the ratio of the pressure of the vapor in the air to the pressure that would be exerted by a saturated vapor at the same temperature.

291. Methods of measuring humidity. 1. The absolute humidity of the air can be obtained by drawing a known volume of air through a series of tubes containing calcium chloride or some other substance which absorbs water vapor. These tubes are weighed before and after the air is drawn through. The increase in the mass of the tubes gives the mass of water vapor absorbed.

ABSOLUTE HUMIDITY AND PRESSURE OF SATURATED WATER VAPOR

Temperature		Grams per Cubic Meter	Pressure (Centimeters of Mercury)	Pressure (Inches of Mercury)
C.	F.			
− 20°	− 4.0°	1.08	0.08	0.03
− 10°	+ 14.0°	2.36	0.20	0.08
− 8°	+ 17.6°	2.74	0.23	0.09
− 6°	21.2°	3.17	0.28	0.11
− 4°	24.8°	3.66	0.33	0.13
− 2°	28.4°	4.21	0.39	0.15
0°	32.0°	4.84	0.46	0.18
2°	35.6°	5.54	0.53	0.21
4°	39.2°	6.33	0.61	0.24
6°	42.8°	7.22	0.70	0.28
8°	46.4°	8.21	0.80	0.32
10°	50.0°	9.33	0.92	0.36
12°	53.6°	10.57	1.05	0.41
14°	57.2°	11.96	1.20	0.47
16°	60.8°	13.50	1.36	0.54
18°	64.4°	15.22	1.55	0.61
20°	68.0°	17.12	1.75	0.69
22°	71.6°	19.22	1.98	0.78
24°	75.2°	21.54	2.23	0.88
26°	78.8°	24.11	2.51	0.99
28°	82.4°	26.93	2.82	1.11
30°	86.0°	30.04	3.17	1.25
32°	89.6°	33.45	3.55	1.40
34°	93.2°	37.18	3.98	1.57
36°	96.8°	41.28	4.44	1.75

2. The dew point is determined experimentally. By the aid of a table similar to the above, the mass of water vapor per cubic meter corresponding to the dew point is obtained. This gives the absolute humidity. The mass required to produce saturation at air temperature is also obtained from the table. The ratio of these two gives the relative humidity.

3. The wet-and-dry-bulb-thermometer method is commonly used in meteorological stations. We have seen in section 280 that a wet cloth hung in the air is cooled by the evaporation from it. The lower the relative humidity, the more rapid the evaporation and

the greater the cooling effect. The bulb of the wet-bulb thermometer is wrapped with a loosely woven cloth, the lower end of which hangs down into a small vessel of water. Capillary action keeps the cloth wet. The *difference* between the readings of a dry-bulb and a wet-bulb thermometer will thus depend on the relative humidity of the air. By an elaborate series of experiments the effect of different relative humidities has been found through a wide range of temperatures. From these values tables have been constructed which are widely used in meteorological stations. Tables are also published in which the dew point is given instead of the relative humidity. Hence, by reading a wet-bulb and a dry-bulb thermometer and using the tables, both the relative humidity and the dew point can be found.

4. Instruments for measuring the relative humidity are called *hygrometers*. The combination of a wet-bulb and a dry-bulb thermometer is one form of hygrometer. A human hair when soaked in ether for twenty-four hours to remove the oil is very sensitive to changes in humidity. Its length alters about one-thirteenth part between a very low and a very high humidity. To make a hygrometer, a long hair is suspended vertically with a small weight at the lower end. Near the lower end the hair is wound once around a small rod or pulley to which is attached a light index. As the hair changes in length owing to changes in humidity, the rod is turned, and the index moves over a scale. This hair hygrometer is a simple and inexpensive instrument, but not a very accurate one. A piece of catgut stretched by a light weight twists or untwists as the humidity changes. Simple hygrometers are made which use this principle.* A wet-bulb and a dry-bulb thermometer, however, give more accurate results than these forms of hygrometers.

292. Fog and cloud. Steam is a clear, transparent vapor, but when it issues into the air it ordinarily changes into a white cloud. This cloud is composed of minute drops of water, the condensed steam. When the vapor comes out into the air, it is cooled and begins to condense as soon as it reaches the temperature where it is saturated. In like manner a cloud is formed when the water vapor in one's breath is cooled below the temperature at which it is saturated (the dew point). All fogs and clouds are produced by a cooling below the saturation point of the water vapor in the air. In general, the distinction between fog and cloud is slight: if it rests on the ground it is called a *fog*; if overhead, it is a *cloud*.

---

*A common form is one where two figures are suspended by the catgut. In dry weather one figure is outside a little house. When the humidity increases, the figure outside goes in and the other comes out.

The small drops of water which form fogs and most clouds fall very slowly. It was stated in section 110 that the air friction for such drops is relatively large. It is only when the drop becomes larger than a diameter of about 0.04 millimeter that the drops fall fast enough to be called rain. Drops 1 millimeter in diameter fall with a speed of about 4 meters per second, and those having a diameter of about 2 millimeters fall with a speed of about 6 meters per second.

When water vapor condenses at a temperature below freezing, snowflakes are formed. When the amount of vapor condensed is sufficiently great, the snowflakes are large enough to fall. When they are tiny, their rate of fall is so small that they apparently float in the air. The high-flying clouds, cirrus and cirro-stratus, consist of snowflakes, or minute ice crystals. The small drops of water composing the clouds are often carried up high into the air where the temperature is below the freezing point. If the drops are small, they may be *undercooled*; that is, cooled below 0° C. without freezing. While in this undercooled state they may suddenly freeze, causing small ice crystals. A haze which sometimes exists high up in the atmosphere and gives a halo around the sun and moon is due to these minute ice crystals.

By an interesting experiment a fog or a cloud may be produced on a small scale on the lecture table. A large inverted flask (Fig. 173) has two glass tubes run up through a rubber stopper. The shorter one of these is connected by a rubber tube a meter or two long to a vessel *B*, which contains water. The longer glass tube runs up nearly to the top of the flask. The other end has a short rubber tube *C*, closed by a pinchcock.

If the cock *C* is closed, and *B* is raised 30 or 50 centimeters higher than the flask, the air in the flask will be compressed by water flowing in. In a closed space of this size, with the relatively large surface of water, the water vapor very soon becomes saturated. If *B* is now quickly lowered to the floor, with the pinchcock *C* still closed, the air in the flask quickly expands. This expansion of the air is a cooling process, for the air is doing work in forcing water out. The result is that *a fog will form in the flask*, for the vapor which was saturated before is now cooled below the saturation temperature. (In order to make the fog clearly visible to a class a narrow beam of light from a projection lantern should pass through the space above the water in the flask.) The fog in the flask will fall very slowly if left alone, for the drops of water are very minute. If the vessel *B* is raised again, the compression warms the air in the flask, and the fog quickly evaporates. The expansion and fog can be produced in another way. If, while *B* is higher than the flask and the air in the flask under compression, the pinchcock *C* is opened, the air in the flask will suddenly expand and the vapor will cool and condense into fog.

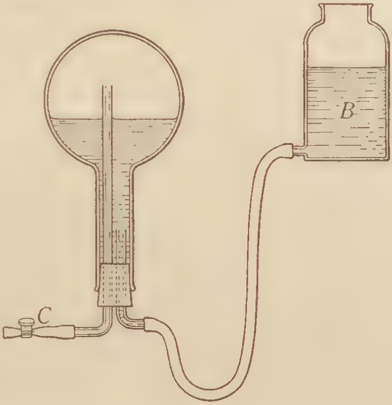


FIG. 173

Another interesting thing about fogs and clouds can be shown by this experiment. Whenever vapor condenses, there must be *something for it to condense on,—a nucleus*. Usually the nucleus of each drop is a dust particle. (It has been shown that electrical ions will also serve as nuclei. But they are present in the atmosphere only in minute quantities except in the neighborhood of a lightning flash and only for a short time afterward.) After the apparatus of Fig. 173 has stood awhile, and especially if the fogs formed in it have settled slowly, most of the dust particles in the flask will have fallen into the water. In that case only faint fogs will be obtained, and sometimes not a trace can be seen. But if, while the reservoir *B* is below the level of the flask, the pinchcock at *C* is opened momentarily, air from the room will rush in.

Usually this air will contain enough dust so that fogs can again be formed, but much denser fogs will be obtained if a little smoke (for example, from a match) is drawn through *C*.

This experiment shows several things of importance:

1. Expansion of air is a cooling process; compression, a heating process.

2. When air containing saturated water vapor is cooled by expansion, a fog will form.

3. In order for the fog to form, dust particles must be present. It is well known that smoke particles assist in the formation of fog. Dense fogs are much more likely to form where there is a large amount of dust or smoke particles in the air. Fogs are formed more frequently in certain cities than in the neighboring country districts. The number of foggy days per year in London has increased with the growth of the city and hence (probably) with the increased dust and smoke in the air.

In small inclosures it is practically impossible for a vapor to be *supersaturated.* But in the open atmosphere, when it is free of dust particles on which the vapor can condense, it is possible for the vapor to become supersaturated; that is, to be cooled below the dew point without condensation of the excess vapor. The following phenomenon was observed on a still, cool morning. The air in a small valley was quiet and clear. A little later a train drawn by a steam locomotive passed up the valley. Almost immediately the valley was filled with a dense fog. Undoubtedly the vapor in the valley had been in a state of supersaturation, but quickly condensed on the dust and smoke particles when they were thrown into the air. The absence in the morning of dust particles in the air of this valley may have been due to the fact that during the night moisture had collected on what particles were present; since there was an insufficient number, the drops of water forming the fog became large enough to cause it to settle rapidly to the ground, thus cleaning the air. Later the temperature was probably lowered still further, leaving the vapor supersaturated.

Ordinarily there are large quantities of dust particles in the atmosphere. In many of our storms there are known to be ascending air currents which can carry dust up to higher altitudes. Some dust is supplied to the atmosphere by the disintegration of shooting stars. Volcanoes have been known to throw dust up very high. In the eruption of Krakatoa, in the East Indies, in 1883, such quantities of dust were thrown up into the higher altitudes that for about three years unusually brilliant red skies were seen at sunset the world over. The dust cloud was carried around the world in about fifteen days by the winds prevailing at high elevations, its progress being traced by the appearance of the red sunsets.

293. Rain or snow. There is a very close relation between the conditions necessary for the formation of rain and those necessary for the formation of cloud. When the amount of vapor that is condensed is small, only very small drops of water may be formed; but if condensation takes place rapidly, some of the drops of water formed will be large enough to fall as rain. When condensation takes place below the freezing point, snow is formed instead of rain. The formation of rain or snow differs from cloud or fog formation in the *rate* and in the *total amount* of vapor which condenses.

It must be clearly understood that rain comes from the condensation of water vapor in the air. Unfortunately, many have learned a nursery tradition that rain comes from the clouds, as if the clouds were large reservoirs of water. Often the water vapor from which a rain is formed will have been in the atmosphere of that section of the country for some time before the rain. Sometimes a condition favorable for rain travels across the entire United States. In the drier sections no rain will fall; but where the humidity is higher, copious rain may be produced. Everyone who has observed weather conditions knows how hard it is for a drought to be broken. Other conditions may be favorable for rain, yet no rain falls because the relative humidity is too low. On the other hand, during a period of excessive rainfall the relative humidity of the air is high, and it is not difficult for rain to form.

294. Conditions causing clouds or rain. The principal conditions which produce a cooling of the air below its dew point and, as a consequence, the formation of fog, cloud, rain, or snow will be discussed under four heads.

1. *The mixing of air currents at different temperatures.* Consider the conditions when warm, moist air—say from the south—is moving close to the ground, while some distance above the ground is a colder wind from a different direction. Between these two moving layers is a continual mixing of the warm and the cold air. If the water vapor in both these air currents is saturated, or nearly so,

there will be at least cloudiness produced. Suppose the temperatures of the two air currents are 10° C. and 22° C. respectively. From the table on page 296 it is seen that if both are saturated, the one at 10° C. will contain 9.33 grams of water vapor per cubic meter and that at 22° will contain 19.22 grams per cubic meter. If equal quantities of air at these two temperatures mix, the resultant temperature will be 16° (assuming that both have the same thermal capacity, which is approximately true), and the water vapor per cubic meter will be approximately 14.27 grams, the average of 9.33 and 19.22. But *saturated* vapor at 16° C. can contain only 13.5 grams per cubic meter. The result is that about 0.8 gram of vapor per cubic meter must condense. However, counter currents of air which are both saturated seldom occur, and the mixing which takes place may produce cloudiness but does not often produce sufficient condensation to cause rain.

2. *Radiation.* Air, especially when it contains considerable water vapor, frequently is cooled by radiation on clear nights. Fogs are often produced by this effect near low ground. But the cooling of air by radiation is not an important cause of clouds and rarely if ever produces rain, for the cooling process is slow.

3. *Warm winds blowing over cold surfaces.* Warm, moist air may be cooled by blowing over cold surfaces. High mountain peaks, especially when covered with snow, are often colder than the general air temperature of the same elevation; hence the warmer wind, blowing against the peak, may be cooled enough to produce the mountain cloud and sometimes rain or snow. In the spring, fogs are often formed by warmer air drifting over thawing ground. The fogs off Newfoundland are caused by warm air currents from the Gulf Stream blowing over colder water. Warm winds blowing from the ocean onto colder land may produce great masses of fog. Especially is this true in the winter, when the land is colder than the ocean.

4. *The cooling of ascending air by expansion.* The most important cause of clouds and rain is the cooling of air by expansion as it ascends. The cooling of dry air, caused by its expansion as it rises, is about 5° F. for each thousand feet. Sometimes there is an uprush of air which may go to an elevation of five miles, where

the pressure is so low that a given mass of air expands to about three times the volume it had at the ground. If there were no water vapor in the air, the cooling effect of this change in elevation would be very great (over 100° F.); but the condensation of the vapor liberates heat (sect. 279), which prevents so great a change.

Air may ascend for reasons of two types: it may be a result of convection, of colder and heavier air forcing up warmer and lighter; or the air may be forced up by a storm or by blowing against a mountain barrier.

The third and fourth conditions often work together. This is especially true in the case of the heavy rains on the Pacific coast from Oregon northward.

One of the most common examples of condensation of the water vapor in an ascending air current is that of a cumulus cloud. It is one of our most beautiful clouds, built up on a fairly flat base, with rounded white masses piled upon each other. In the daytime the air near the ground gets heated, and here and there rising air columns carry up this warmer air. Each of these clouds is the head of one of these convective currents. When the air reaches such a height that expansion has cooled it to its dew point, condensation begins. Since this altitude is quite definite, the bottoms of these clouds are fairly flat. The air current rising up through the cloud blows it up and out into the beautiful shapes it takes. There are well-known instances of a cumulus cloud hanging almost continuously on fair days over a small island. The air over the island, being warmer in the daytime than that over the surrounding water, is forced upward, causing a nearly permanent column of rising air topped with a cumulus cloud.

Those who are familiar with the mountains or with photographs of them have noted the frequent occurrence of clouds hanging on the mountain sides. The heat from the sun falling on the mountain slopes warms the ground and the adjacent layers of air. This warmer air flows up the mountain side, and is cooled by its expansion until its temperature is lowered below its dew point, with a condensation of part of its vapor into a cloud.

At the equator, where the general convective motions of the earth's atmosphere cause ascending currents, there are frequent heavy rains. The heavy rains of our thunderstorms are caused by an unusually rapid uprush of warm moist air. The air in that general disturbance called a *low*, or *cyclone*, is ascending; hence a low usually brings cloudiness and rain. On the other hand, in the disturbances called *highs*, or *anticyclones*, air is descending. As air descends, it is warmed by compression, and its relative humidity decreases. Hence in these highs the skies are clear and the winds relatively dry (see section 228).

When the prevailing winds are in such a direction as to add to the sea breeze, which is the convection current produced by the fact that the land is warmer than the water, large quantities of warm moist air are carried inland and upward by the rising current of the interior. The eastern coasts of Florida and Mexico receive the trade winds and, as a result, have a large rainfall.

In the United States, on the average, the direction of the wind is from the west. As it blows inland from the Pacific coast it is deflected upward by the mountain ranges and descends on the eastern slopes. After crossing the mountains the general slope is downward, the plains of eastern Colorado being over four thousand feet higher than the Mississippi River. The result is that the coast receives generous rains, while in the interior the annual rainfall is small until the Mississippi Valley is reached. If the prevailing winds were from the east, the eastern slopes of the Rockies would receive a far greater rainfall, and the Pacific coast would be relatively dry.

If the dew point of the air at the surface is known, the height to which air must rise before cloud formation begins may be computed as follows: Air is cooled about 5° F. for each thousand feet of rise. As the air rises it expands, thus decreasing the vapor density and lowering the dew point (the dew point is lowered about 1° F. for each thousand feet of ascent). Owing to these two facts the air as it rises *approaches its dew point* at the rate of approximately 4° F. for each thousand feet. If the air temperature is 85° F. and the dew point 70° F. (corresponding to a relative humidity of about 63 per cent), the temperature of the air would have to be lowered 15° F. to reach its dew point. Or the dew point would be reached by an ascent of

$$\frac{85-70}{4} = 3\tfrac{3}{4} \text{ thousand feet.}$$

A result obtained in this way must be regarded as only an approximation.

295. Winds: general causes. In section 242 a short account was given of the production of convection currents in the earth's atmosphere owing to the unequal heating of the earth's surface. Brief accounts were also given there of the so-called land and sea breezes and of the mountain and valley breezes. In this and the following sections the circulation of the air will be more fully explained; but the motion is so complicated, occurs on such different scales, and is so modified by local conditions that this discussion must be regarded as somewhat superficial.

Convection currents in a gas or liquid are produced when the gas or liquid is heated at the bottom or cooled at the top. Both these conditions exist in the atmosphere, as shown by the following considerations: Much more than half of the radiant energy received from the sun is absorbed by the atmosphere; but most of this absorption takes place in the lower layers, for two reasons. The first of these is that the atmosphere has its greatest density at the lower levels, considerably over half of the total mass being within less than four miles of the surface of the earth. The other is that much of the ability of the air to absorb radiant heat is due to the water vapor it contains, and the amount of water vapor in the air is much greater near the ground. Furthermore, the lower layers of air receive heat from the ground by radiation and conduction. Convection is further aided by the fact that the air at high altitudes is cooled by radiation.

The largest convection currents taking place are due to the great amount of heat the earth receives in the equatorial regions. At the equator the air rises, flowing north and south. On each side of the equatorial zone the surface winds blow toward the equator. These are called the *trade winds*, while the higher currents, flowing away from the equator, are called the *anti-trade winds*.

Half of the earth's surface lies between the parallels of 30° N. and 30° S.; that is, south of New Orleans and north of a corresponding point in the Southern Hemisphere. Hence approximately half of the earth's atmosphere lies between these parallels. One of the consequences is that most of this general circulation lies within that region. Part of the air which rises at the equator descends in belts parallel to the equator: that in the Northern Hemisphere is known as the calms of Cancer and that in the Southern Hemisphere as the calms of Capricorn. Sometimes these regions are called *horse latitudes*. The descending currents which turn back toward the equator give the trade winds. Hence the trade winds lie between these calm belts and the equatorial belt of ascending air. The equatorial belt is also a belt of calms, often called the *doldrums*. The trade and anti-trade winds are interfered with by storms and by local conditions; for example, by mountain ranges and by the unequal heating of land and sea. But on the ocean away from the land there is less interference. There the trade wind has been known to blow continuously for weeks. When most of the world's trading was carried on by sailing-ships, the belts of calm and wind were of great importance in commerce.

It has been pointed out that the usual condition for rain is an ascending air current and that descending currents are accompanied by clear skies. This is well shown in these belts of calm. At the equator, or doldrums, the air is ascending, and the rainfall is very heavy; on the other hand, in the horse latitudes the direction of the general air current is downward, and there is little rain. Most of the arid regions of the world are in the horse latitudes; but in some places (for example, in the Gulf states) this general condition is changed by other factors.

In general, we may say that the chief cause for wind is convection, the hotter and lighter air being forced up by the inflow of colder air from elsewhere. In some types of storms (thunderstorms, for example) much of the energy comes from the heat liberated by the condensation of water vapor. Some energy is received from the rotational energy of the earth, as explained later. But the chief source of the energy of the winds, including the violent storms which are so common in many parts of the world, is the sun.

296. Effect of the rotation of the earth. In section 131 it was explained that on account of the rotation of the earth a projectile is deflected toward the right. It was also pointed out that the same effect is produced in the winds. Air flowing southward drifts to the *west* (that is, to the *right* as one looks in the direction it is blowing). For this reason the trade winds do not blow directly southward in the Northern Hemisphere, but toward the southwest. The anti-trade winds, those blowing away from the equator, drift eastward; that is, to the *right*. This tendency for the wind to be deflected to the right is true for east and west winds as well as for north and south ones.

There is a very important air motion extending over all regions of the earth except the equatorial ones; namely, the prevailing direction of the winds from the west. This is not so noticeable in the eastern half of the United States, for storms follow each other so closely that the *surface* winds blow from every direction. Yet the average direction is from the west. At high elevations the effect is quite pronounced. On Pikes Peak approximately 70 per cent of the winds come from the west. At higher elevations the western wind is very constant. This can very easily be seen by watching the *cirrus* clouds, those delicate wisps which fly at an average height of about five miles. Rarely indeed do these move in any direction except eastward. This west wind is called the *prevailing westerly.*

The prevailing westerly is due in a large part to the rotation of the earth. The rising air at the equator divides and flows out toward the poles. Let us consider the case for the Northern Hemisphere, part of which is represented in the circumpolar diagram of Fig. 174. The wind flowing northward from the equator is deflected to the right, as shown on the edge of the diagram. This sets up around the pole a whirling motion which extends through the temperate zone.

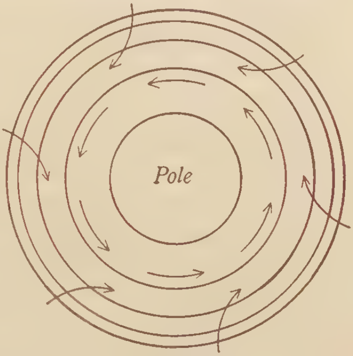


FIG. 174

A more rigorous mode of looking at this is as follows: The air at the equator, having the motion of the earth, has a very large amount of angular momentum (sect. 134). As it moves northward its radius of rotation is decreased, and in order to conserve its angular momentum the speed must increase (sect. 135). If it were not for frictional losses the speed acquired would be enormous. It has been estimated that it would reach hundreds of thousands of miles per hour. In the Southern Hemisphere, where there is less land, the speed of this westerly at the surface is much greater, and the wind often becomes a gale.

297. Cyclones. The term *cyclone* is applied to those large whirlwinds in the atmosphere which have diameters ranging from a few hundred to several thousand miles. It should not be applied to our *tornadoes*, which are smaller but more destructive. In a cyclone the winds blow toward and around the center in a spiral motion, counterclockwise in the Northern Hemisphere. These huge whirls are carried eastward in the temperate zones by the prevailing motion of the air, in a manner similar to the way eddies in water are carried along by a stream. The eastward drift is not uniform: some cyclones move fifteen hundred or more miles a day, while others may move only about a hundred miles in that time. In that part of the United States and Canada lying east of the Rocky Mountains these cyclones are very common. A daily weather map which does not show the presence of at least one of these cyclones is rare. The continuous changes in the weather of the eastern half of the United States and Canada at all times of the year are due to the approach and passage of these cyclones.

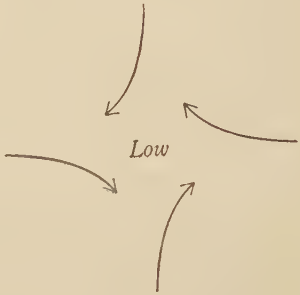


FIG. 175

The center of the cyclone is a large ascending column of air. Since this air is flowing up, there must be winds blowing in toward the center along the surface of the ground. But, as we have seen before, the rotation of the earth causes these to deflect to the right (Fig. 175). This produces the spiral motion which is characteristic of a cyclone.

The barometric pressure at the center of a cyclone is below normal. The reasons for this are simple. The ascending mass of air in the center is warmer than the air of the same elevation at distant points; it therefore has a smaller density and weight. Hence the pressure it exerts must be less. There is also a difference in pressure between the outside and the center on account of the rotational motion; for in order to force air toward the center the difference in pressure must be greater than the centrifugal force. The low barometer at the center is so marked that these cyclones are usually referred to by weather forecasters as *lows*. The barometric pressure at the center will average about 29.6 inches (corrected for elevation), and it may in extreme cases be below 28 inches.

Fig. 176 shows the *isobars*, or lines of equal barometric pressure, and the wind directions in a cyclone having a diameter of perhaps twelve hundred miles.

There is considerable doubt as to the causes which are the most important in the starting of these whirling motions. In the tropics and in many of our summer lows the cause is probably convection. When warm, moist air rises, the air cools by expansion until some of the water vapor condenses. This process of condensation releases a large amount of heat; for the heat of vaporization of water is very large, and the air becomes warmer than it ordinarily would be at that elevation.* This increases the convective effect, and the upward movement increases until a large mass of air is involved. The surface winds flowing in produce the whirling motion, and the prevailing motion of the air carries the whirl along. It has, however, been shown that the foregoing does not explain the majority of the lows of the United States and Canada. It is found that these lows can start anywhere. They can form in the winter time in the frozen Northwest, where the air is very cold and dry. Explanations have been suggested, but it is generally conceded that these are incomplete.

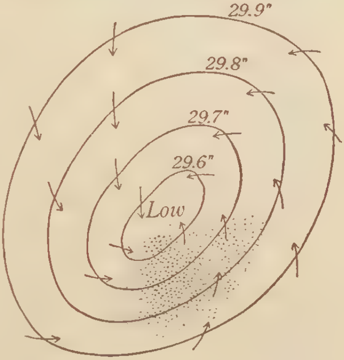


FIG. 176. The isobars and wind directions of a cyclone

298. Anticyclones. While the cyclone is a region of *ascending* air currents, the *anticyclone* is an extensive region of *descending* air. It is obvious that there must be in the circulation of the air as much air descending as ascending. Since the center consists of a descending mass of colder and therefore heavier air, the air pressure at the center is above normal. For this reason anticyclones are commonly referred to as *highs*.

Since the air is descending, the surface winds must flow out from the center. These surface winds are deflected to the *right*, in the Northern Hemisphere, by the rotation of the earth.

---

* The heat liberated by condensation of 1 inch of rainfall is greater than would be received from the sun in several days.

This produces the spiral motion in a clockwise direction which is characteristic of an anticyclone (Fig. 177).

It is thus seen that the anticyclone, as its name indicates, is the opposite of the cyclone: the air is descending instead of ascending, and the spiral rotation is in the opposite direction. Moreover, as will be explained in the next section, the weather conditions produced are opposite in character.

The anticyclones, or highs, are also carried along by the prevailing westerly, so that their general motion is eastward. While many of the lows traveling across the eastern half of the United States originate in this country, the highs which sweep across come from beyond our boundaries.

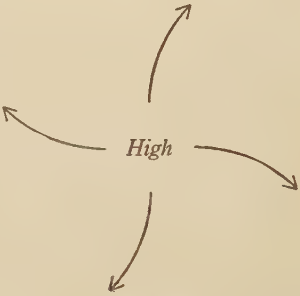


Fig. 177

299. Effect of cyclones and anticyclones on the weather. It has been shown that an ascending current of air produces cloudiness and often rain or snow, while descending currents are accompanied by clear skies. Hence, as a low with its ascending air currents approaches, cloudiness increases, and usually we have rain or snow; but as the high which often follows the low approaches, the barometer rises, the skies clear, and (since the descending air is warmed by compression and thus becomes relatively dry) the humidity diminishes.

In front of a low the wind is from the south or southeast and is likely to be warm and humid. In the winter we say that the weather is moderating; in summer it may become hot and sultry. In the southeastern quadrant of a low the wind at the surface of the ground is from the south and is usually warm and humid. Usually it blows in under a mass of cold air above, a condition which tends toward instability. Hence the southeastern part of the cyclone* has not only most of the rainfall but also the thunderstorms and sometimes tornadoes.

---

*The shaded part of Fig. 176.

As an anticyclone which travels north of us approaches, the wind shifts into the northwest, becoming cool and dry. In the winter the sudden change of temperature which gives us a cold wave is due to a high coming from the west or northwest, closely following a low. When the highs move through the Gulf states, the effect on the weather of the country to the north is very different. This is quite marked in summer, when the highs and lows may move eastward rather slowly. When a low moves slowly across the northern part, with the high moving across the southern border, we may have a hot wave. For days there may be a surface wind blowing up across the Southern states. A succession of these highs moving south of us will give an almost unbroken period of hot weather. But when the summer highs travel north of us, and the lows to the south, we have unseasonably cool weather. (These statements apply to the part of the United States lying east of the Rocky Mountains.)

The daily weather maps show the position of the highs and lows; and since their probable motion is known, the weather can be forecast for any locality.

300. Thunderstorms. Thunderstorms, especially the severe ones, usually occur in the late afternoon of a hot, sultry, and oppressive day,—a day when dew will form on the outside of a glass containing only moderately cold water. During the day there has usually been a light breeze from the south or southeast, but the air feels stagnant. The storm usually appears in the west and moves eastward, often against a light southeast breeze. The typical thunderstorms usually occur in the south or southeastern section of a low, where the surface wind is a warm, humid layer of air blowing from the south or southeast. This warm air is usually forced in under a large mass of cold air, developing an unstable condition, so that here and there the warm lower layer breaks through and starts a convection current.

There are many types of thunderstorms, but they all have the same cause. They are caused by the rising of warm, moist air. As this air rises it is cooled by expansion until the dew point is reached, when condensation begins. In a well-developed storm the uprush of air is so great and carried to such a height that there is a very rapid rate of condensation, and the rainfall may be very heavy. The storm usually extends up into the prevailing westerly of the higher altitudes, so that it is carried along by this eastward drift of the atmosphere.

The motion of the air currents in the typical thunderstorm is of special interest. Fig. 178* shows a section of one, with the direction of the wind indicated by arrows. The storm is moving to the right. In front of it the warm, moist air is moving toward and up into the storm. This current supplies the water vapor which is condensed. The descending currents are cold air from above, which takes the place of the rising currents. To be more exact, it is this heavy, colder air that forces up the warmer and lighter air in front. The sudden drop in temperature as the thunderstorm approaches, caused by this cold air from above, is always well marked. The upward rush of air in front and the descending air in the storm cause the whirl, or rotational motion, about a horizontal axis which is typical of the front part of our summer thunderstorms.

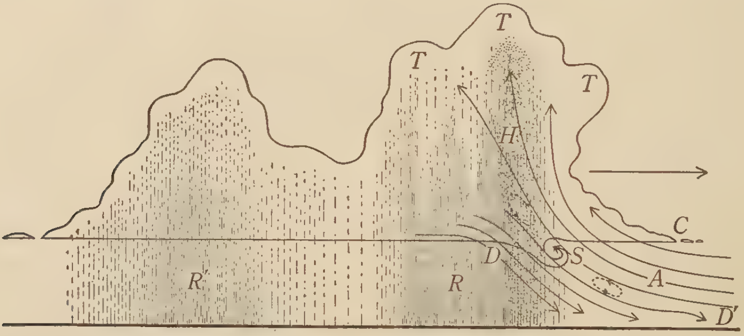


Fig. 178. Ideal section of a typical thunderstorm. *A*, ascending air; *D*, descending air; *C*, storm collar; *S*, roll scud; *D′*, wind gust; *H*, hail; *T*, thunderheads; *R*, primary rain; *R′*, secondary rain

The heat liberated by the condensation of so much water vapor furnishes much of the energy of such a storm. This heat warms the uprushing currents, so that even at a high elevation the air is warmer than that on either side of it. This increases the speed at which it rushes up and also the speed of the downward-moving cold air. Sometimes this upward-rushing air carries with it drops of water to such an elevation that they become frozen. They then fall as *hail*. Rain cannot fall through air with a relative velocity greater than 8 meters per second (sect. 110); hence if the upward current of air has a velocity greater than that, the drops are all carried upward. (It is now known that the intense electrification produced in such a storm is due largely to the way in which the uprushing air breaks the drops of water into a spray. Electricity is generated by such a process.)

---

* This figure is taken from Humphrey's "Physics of the Air."

301. Tornadoes. Rotational motions of the air are formed on all sorts of scales. They vary in size from the small dust whirl, or whirlwind, so commonly seen, to large, cyclonic motions several thousand miles in diameter. The small dust whirl is usually started by a convection current which forms on a sunshiny day over a hot roadway or bare ground. The air which rushes in is deflected by obstacles, so that a rotary motion is given. The direction of this rotation is purely accidental; but in the larger storms the direction of the rotation about a vertical axis in the Northern Hemisphere is always counterclockwise.

The relatively small rotary storms which have so great a destructive effect in the United States are *tornadoes*. They have a very well-developed counterclockwise rotation about a vertical axis, and always occur in the southern or southeastern part of a cyclone and from two hundred to eight hundred miles from its center, under conditions which are favorable for severe thunderstorms. Apparently the cause is a very violent upward convection current. The air blowing in from the sides is deflected, either by the rotation of the earth or, more probably, by the general cyclonic motion in which it is an eddy, so that a counterclockwise whirling motion is produced. On account of the violent uprush of the air in the center, the whirling mass of air moves toward its center rapidly. It has been explained before that when air or any mass which has angular momentum has its radius of rotation decreased, there must be an increase in the speed. It behaves like water running out of an opening in the bottom of a washbowl (sect. 135).

Certainly the wind in a tornado does attain enormous speeds. This speed has never been measured, but estimates indicate that it may be as high as five hundred miles per hour. On account of the enormous speed of rotation the centrifugal forces are very great, and atmospheric pressure on the outside of the tornado must be very much larger than that at the center. The pressure at the center of a severe tornado has never been measured, but it must be quite low.

The destructive effect is due not only to the high speed but also to the sudden drop in the barometric pressure, causing the air inclosed in buildings to expand outward with explosive violence. Windows and sometimes walls are forced outward.

The path of a tornado is from a few feet to a thousand yards wide and from a mile to several hundred miles long. Tornadoes occur almost exclusively in the part of the United States which lies east of the Rocky Mountains and are most frequent over level country. In hilly or mountainous country the topography encourages the flow of convection currents, and this is probably the reason that the excessively unstable conditions which sometimes occur over a level country are avoided.

302. The averaging of temperature. It is found that the daily temperature of any one place, when averaged over twelve months, is nearly a constant quantity, varying only a few degrees from one year to another. It does not make any difference what are the limits of the period taken, whether from January to January or from June to June. For example, a very hot August can be averaged in with the preceding eleven months or with the succeeding ones. But because this average is approximately constant, many argue that the weather must change in such a way that this average will be a constant; for example, a hot summer must be followed by a cold winter. Does the student think this reasoning sound?

The temperature of the earth's surface is determined by the rate at which it receives heat from the sun and by the rate at which it loses heat into space. Since both these rates, when averaged over the surface of the earth, are practically constant, it follows that the temperature at any one time, averaged over the entire surface of the earth, must be a constant.* If at any place the temperature is unusually high, it is because it is receiving more than its share of the warm winds. Some other place must be getting the cold winds. Indeed, the facts show that when one locality has a hot season, another has a cold one. *The average of temperature is struck at the time*, not months later.

A study of the factors that cause changes in the weather shows that when it is unusually cool in one locality it should be warmer in some other. For example, a high traveling eastward should bring warmer weather to the regions north of its track and cooler weather to those lying to the south.

303. The averaging of rainfall. There is a relatively small amount of water stored in the form of vapor in the atmosphere. Estimates show that if all the vapor were condensed and fell as rain, it would give less than two inches of water. In other words, the storage capacity of the atmosphere is relatively small, and water must fall as rain at almost the same rate at which it evaporates. The greater part of the earth's surface is water, and the total evaporation from this is fairly uniform; hence the total daily amount of water falling as rain must be approximately constant. From this it follows that if any section of the country is having a rainfall above the average, there must be other localities which are not receiving their share. There can be no such thing as a world-wide drought nor a world-wide excess of rain.

PROBLEMS

1. The dew point was found to be 12° C. when the air temperature was 20°. Find, by the use of the tables in the text, the absolute and relative humidities; also the pressure of water vapor.

2. On a humid day in summer the dew point was 78.8° F., and the barometer 74 cm. What part of the atmospheric pressure was due to the water vapor?

3. The relative humidity on a day when the temperature was 30° C. was 63 per cent. Compute the mass per cubic meter of water in the air.

4. The dew point was found to be 68° F. on a day when the temperature was 86° F. Find both the absolute and relative humidities.

---

*Any appreciable change in the average temperature of the earth would mean a tremendous change in its thermal energy.

5. Air temperature is 30° C., and the relative humidity 60 per cent. Estimate the value of the dew point.

6. How many grams of water must be evaporated into a room 10 m. square and 4 m. high to raise the relative humidity from 30 per cent to 70 per cent when the temperature is 20° C.?

7. The dew point near the ground was 24° C. The air rose until its volume was doubled, and it became saturated at 0° C. Compute the number of grams of water vapor condensed from each cubic meter of the air, measured at the ground.

8. A cubic meter of air has at normal pressure a thermal capacity of about 250 cal. How much would its temperature be increased if it received all the heat liberated by the condensation of 10 gm. of water vapor? (Assume the heat of vaporization to be 600 cal./gm.)

CHAPTER XXIII

HEAT ENGINES

Heat engines, 304. The reciprocating steam engine, 305. Efficiency of an engine, 306. Steam turbines, 307. Maximum theoretical efficiency, 308. Internal-combustion engines, 309.

304. Heat engines. The earliest heat engine was probably the toy invented by Hero, a pupil of Archimedes. Steam was generated in a hollow ball which was partly filled with water. This steam escaped through two tubes on opposite sides of the ball. The ends of the tubes were bent so that the reaction from the steam jets supplied the motive force to set the ball in rotation. This reaction principle is today used in some types of the steam turbine.

There are many types of machines now in use for converting heat energy into mechanical energy. The great development of these has been due in part to the large supplies of fuel from which energy in the form of heat can be conveniently obtained. In several types of engines the fuel is used to generate steam, and the energy in the steam is used to drive the engine, as was the case in Hero's toy. In another class the fuel is actually burned inside the engine. The commonly used gasoline engine is one type of these internal-combustion engines. There are other special types (the hot-air engine, for example), but space will not permit an explanation here of all the various types which have been developed.

305. The reciprocating steam engine. The development of the modern steam engine has been one of gradual evolution, many men making important contributions; but the work of James Watt (1736–1819) was so important that he is often given the credit for its invention. Watt's engine was far more efficient than any of the earlier types, requiring for the development of the same mechanical energy only about one fourth as much coal as its predecessor.

An understanding of the way steam under pressure can be used to run an engine may be obtained by aid of the simplified diagram of Fig. 179. The piston *P* is connected by the driving-rod *R* to the shaft *S* by a crank, as shown. The eccentric rod *R′* is connected to the slide valve *B*, which, as it slides to and fro, first uncovers the passageway *N* into the right end of the cylinder and then, closing it, uncovers *M*, leading into the left end. When steam under pressure enters at *A*, it flows through *N* into the right end of the cylinder, forcing the piston *P* toward the left. At this stage the steam on the left side of the piston can escape through *M* and out of the exhaust *E*. As the piston moves, it turns the shaft *S* by means of the driving-rod and crank. This causes the slide valve *B* to move to the right. The slide valve thus first cuts off the supply of steam through *N*, and then a little later closes the exhaust from the left end of the cylinder. When the piston has reached the far end, the right end of the cylinder is connected through *N* to the exhaust *E*, and live steam is admitted to the left end of the cylinder through *M*. Thus, as the piston *P* moves back and forth, the slide valve also moves back and forth, first admitting steam into one

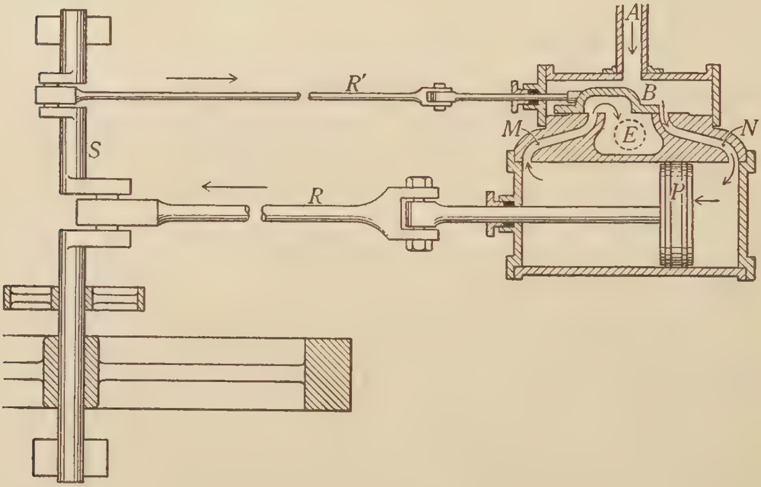


FIG. 179. The steam engine

side of the cylinder, then at the proper time turning the steam into the other, while the first side is connected to the exhaust, and so on. Thus a continuous reciprocating motion is given to the piston, and a continuous rotation to the shaft. In a stationary engine a heavy flywheel is necessary to insure the passing of dead centers and to keep the motion uniform.

More energy can be extracted from the same amount of steam if the supply to one end of the cylinder is cut off before the piston has reached the end of its motion. Before the steam supply is cut off, the steam in the cylinder is at high pressure and temperature, nearly that of the boiler. After the valve is closed, the steam in the cylinder expands, driving the piston on. In this way the vapor does work and loses energy; hence its temperature and pressure fall. If the valve is properly adjusted, the pressure of the steam will fall to nearly that of the place into which it passes from the exhaust. If the steam had been kept at high pressure in the cylinder, the work it could do by its own expansion would have been lost. The lower the temperature at which the steam is released, the more heat energy of the steam has been utilized.

If the steam escapes into the atmosphere, the engine is called *noncondensing*; but if the steam is led into some place where it is condensed at a low temperature and pressure, the engine is called a *condensing* engine. The *condenser* is an air-tight chamber the interior of which is maintained much below the atmospheric pressure. As the steam enters this chamber it is condensed by a spray of cold water. If the temperature of the condenser is very low, the pressure of the vapor in it cannot be higher than the pressure of water vapor saturated at that temperature. It should be obvious that when the steam is forced out from the cylinder into the atmosphere, it exerts a back pressure on the piston of about 15 pounds per square inch. But in a high-grade condensing engine the pressure in the condenser may be as low as 1 pound per square inch, or about 14 pounds per square inch lower than atmospheric pressure. When the steam exhausts into such a condenser, the back pressure on the piston is very small. Hence the steam driving the piston forward can expand to a lower pressure and temperature, thus giving up more of its heat energy.

In the *compound* engine the steam passes through several cylinders in succession. In the first cylinder the steam expands to an intermediate pressure. At this reduced pressure the steam goes into the second cylinder, and from there it may exhaust into the condenser. In the triple-expansion engine the steam passes through three cylinders, the pressure being reduced in each one. Quadruple-expansion engines are now in use, where the same steam passes in turn through four cylinders. The most effective steam engines today are quadruple-expansion condensing engines.

It is important for the student to understand the general transfers of energy that take place in a steam-power plant. In the first place, the fuel has a definite amount of energy. When the fuel is burned, the energy is converted into heat energy. Part of this heat energy is used to raise water to its boiling point and to convert it into steam under pressure. The steam under pressure does work in two steps. First, when steam is forced into the cylinder, the steam in the boiler does work in expanding. This is a cooling process, and the steam in the boiler loses energy. Then, when the steam in the cylinder is cut off from the boiler, it expands, doing work, and loses heat energy as its pressure and temperature fall. In both steps expanding steam does work on the piston. In this way mechanical work is done at the expense of heat energy.

306. Efficiency of an engine. There are several different quantities which are called the *efficiency* of an engine. Inasmuch as all these are commonly used, it is important to know the differences between them.

The *thermal efficiency* may be defined as the *ratio of mechanical work per hour to the available energy in steam supplied per hour* (the two must be measured in the same units). The "available" energy of the steam is the difference between the heat energy the steam has at the high temperature of the boiler and the heat energy it has after it has condensed into water. It may be stated as the amount of energy required to put the water from the condenser back into steam in the boiler. The "mechanical work" usually includes the total mechanical energy developed, including that lost in the engine on account of the friction of its bearings. The *rate* at which this energy is developed is often referred to as the *indicated horse power*. But the rate at which the shaft of the engine supplies energy to other machinery is called the *brake horse power*, because it is what would be measured by some sort of dynamometer or Prony brake. If, instead of the total energy developed, the delivered energy is used in computing the efficiency, the result is called the *overall efficiency*. Thermal efficiencies of about 25 per cent have been obtained with multiple-expansion engines.

In a steam plant the efficiency of the boiler is an important item. The *boiler efficiency* is defined as the *ratio of the heat energy in steam generated per hour to the energy in fuel used per hour*. In good plants this will usually run about 60 per cent, although 85 per cent has been reached.

It is often convenient to group the engine and boiler together as one unit and define the efficiency of the combination as the ratio of the mechanical work per hour to the energy in fuel used per hour. Instead of stating this as a percentage, it is quite common to state this efficiency by the number of *pounds of coal required for each horse-power hour* or *for each kilowatt-hour*. An efficiency of about 1 pound of coal per horse-power hour has been reached; some modern plants develop 1 kilowatt-hour for each $1\frac{3}{4}$ pounds of coal used.

307. Steam turbines. The steam turbine has in recent years come into common use in power plants where large or moderately large units are needed. The simplest form is the *impulse*, or *velocity*, type. Steam issues from one or more nozzles at a very high speed, from 1000 to 4000 feet per second. These high-speed jets strike the vanes of a wheel and set it in rotation, in much the same manner as water turbines are given energy by jets of water. The wheel acquires a very high speed, which is most efficient when the vanes have about one half the speed of the jet. In this form of engine the energy of the steam in the boiler is changed into kinetic energy of the jet, and this in turn gives up energy to the moving vanes.

Many of the more modern types are much more complicated. In these the steam, instead of losing its pressure at one place, passes through a number of stages, in each of which it loses part of its initial pressure, acquiring kinetic energy and then acting on moving blades. Since the steam expands more gradually in this type of engine, it does not acquire such a high velocity, and the turbine does not need to have the very high velocity which the simple impulse type must have.

For the same horse power a turbine occupies less space than a reciprocating engine. It has, also, greater freedom from vibration. It is not so efficient as the best reciprocating engine when high-pressure steam is used, but it is more efficient than an engine for low-pressure steam. In large sizes the turbine costs less than a reciprocating engine of the same horse power.

To run electrical generators large turbines of from 20,000 to 30,000 horse power have been constructed which give an efficiency, from fuel to the electrical energy developed, of about 23 per cent, a result close to that of the best internal-combustion engine.

308. Maximum theoretical efficiency. In the subject of *thermodynamics* two important theorems are proved for the *Carnot* engine. This is not an actual engine, but an imaginary one in which all processes are carried on under ideal conditions, without losses of heat energy by radiation, conduction, etc. The "working substance" receives all its heat energy at a certain high temperature T_1; and when it does work in the engine, it is cooled down to the temperature T_2.

The first of these theorems is as follows: *No engine can have an efficiency greater than that of the Carnot engine when working through the same range of temperature.*

The second theorem is stated thus: *The thermal efficiency of a Carnot engine is equal to*

$$\frac{T_1 - T_2}{T_1},$$

where T_1 and T_2 are the high and low temperatures, in the *absolute* scale, between which the engine works. It follows from these two theorems that *no* engine can have an efficiency greater than this fraction. For example, suppose that a given steam engine received steam from a boiler the temperature of which is 200° C. (or 473° K.), and delivers this steam to a condenser at a temperature of 40° C. (or 313° K.). The thermal efficiency of a Carnot engine working between these limits is

$$\frac{473 - 313}{473} = 0.34.$$

Hence an actual engine working between these temperatures cannot have an efficiency greater than 34 per cent. The actual efficiency attained for these temperatures is about 25 per cent.

This formula for the efficiency shows that an engine can be made more efficient by increasing the difference between the high and the low temperature. The high temperature is often increased by superheating the steam; that is, by heating it further after it has left the boiler. Lower temperatures are obtained by producing better vacua in the condensers, thereby condensing the steam at a lower temperature.

309. Internal-combustion engines. The most commonly used type of *internal-combustion* engine is the *four-cycle gas engine.*

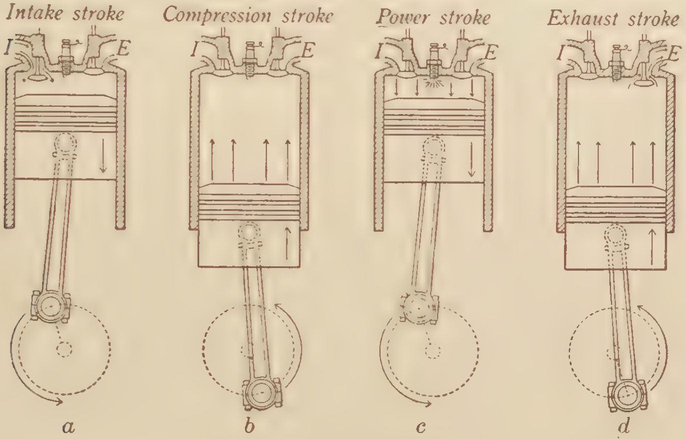


FIG. 180. The gas engine

In Fig. 180 are shown the four stages in the operation of a single-cylinder engine. For simplicity many details are omitted.

In Fig. 180, *a*, is shown the *intake* stroke; the explosive mixture is being drawn into the cylinder. The intake valve *I* is open; the exhaust *E* is closed. The *compression* stroke is shown in Fig. 180, *b*. This motion compresses the gas into the end of the cylinder, where, by an electric spark, the gas is exploded, driving the piston forward. This stage is shown in Fig. 180, *c*. The *exhaust* stroke (Fig. 180, *d*) is for the purpose of driving out of the cylinder the gases which result from the combustion. After this is done the operation is repeated. The axle of the engine makes two complete revolutions for each explosion. In order to keep the engine running between explosions, a heavy flywheel is mounted on the axle.

In many engines more than one cylinder is used. Thus the four-cylinder engine is one in which four cylinders are connected to the same axle. These are usually so set that there is one explosion for each half-revolution of the axle. The best four-cycle gas engine is able to convert about 30 per cent of the fuel energy into mechanical energy.

The *two-cycle* engine explodes once for each revolution. There is not a separate stroke used for the exhaust, but the explosive mixture rushing in under pressure drives out the exploded gases.

The *Diesel oil* engine differs from the gas engine in many respects. In this engine air is compressed in the cylinder by a return stroke of the piston until heated to a temperature higher than the firing point of the fuel. The oil under pressure is then blown in the form of a fine spray into this compressed air. At the high temperature of the air the oil begins at once to burn. There is no sudden explosion, but the oil burns as admitted. The gases resulting from this combustion exert, as in the case of the ordinary gas engine, large pressures; for at ordinary pressures they occupy a much greater volume than the materials which were burned. Formerly, Diesel engines were made only in stationary units and not in small sizes. Their use in locomotives, automobiles, and the like is still in an experimental stage. Large Diesel engines may convert 40 per cent of the fuel energy into mechanical energy.

PROBLEMS

1. The average resultant pressure on the piston of a steam engine was 60 lb./sq. in., the length of the stroke 10 in., the area of the piston 60 sq. in., and the number of revolutions per minute 200. Compute the horse power.

2. How many pounds of coal per horse-power hour will be needed for a plant which has an efficiency of 15 per cent, if the coal used has a heat of combustion of 14,000 B.T.U./lb.?

3. A large steam plant used 1.5 lb. of coal for each horse-power hour, burning coal which had a heat of combustion of 14,500 B.T.U./lb. What percentage of the total energy was obtained?

4. The heat of combustion of gasoline is 20,000 B.T.U./lb. If the efficiency of the engine is 20 per cent, how many horse-power hours can be developed from each pound of gasoline?

PART III. WAVE-MOTION AND SOUND

CHAPTER XXIV

WAVE-MOTION AND SOUND

Wave-motion, 310. Transverse waves, 311. Longitudinal waves, 312. Transfer of energy in a wave, 313. Water-waves, 314. Relation between velocity, period, frequency, and wave-length, 315. Standing waves, 316. Sound, 317. Sound-waves longitudinal, 318. Speed of sound, 319. Determination of the direction of a source of sound, 320. Sound-ranging, 321. Reflection of sound-waves, 322. Refraction of sound-waves, 323. Musical sounds, 324. Pitch, 325. Intensity, or loudness, 326. Tone quality, 327. Resonance, or sympathetic vibration, 328. Interference, 329. Beats, 330. Doppler's principle, 331. Vibrating strings, 332. Frequency of vibration of strings, 333. Vibrating rods, 334. The Kundt's-tube experiment, 335. Vibrations of air columns, 336. Musical intervals, 337. Law of harmonious combinations, 338. The major scale, 339.

310. Wave-motion. The study of wave-motion is connected, in this text, with that of sound, not because sound is the only form of wave-motion but because there is some advantage in treating them together. It is easier to understand some of the properties of waves in connection with a definite type of wave than by studying them in the abstract. But this coupling of the two subjects must not be permitted to deceive the student. There are many other important applications of wave-motions. The vibrations, or tremors, of buildings, bridges, and other structures are ordinarily wave-motions. In the subject of light, frequent reference will be made to the principles explained here. Telephonic transmission along wires and cables is a special type of wave-motion. Electrical waves in space are now used in wireless telegraphy and telephony. In almost every field of physical science one finds instances of wave-motion.

Waves may be classified according to the direction of the vibration of the medium as compared with the direction the wave is moving. According to this mode of classification there are two types: (1) *transverse* waves, those in which the direction of vibration is at right angles to the direction of propagation of the wave; (2) *longitudinal* waves, those in which the direction of vibration is parallel to the direction of propagation. (In a few cases (water-waves, for example) the motion is a combination of these two types.)

311. Transverse waves. Imagine a stretched cord with waves traveling along it. Fig. 181 shows what might be termed an instantaneous picture of it. Each particle of the rope is vibrating up and down with the same frequency. At *a*, *e*, and *i* the particles have reached the highest points of their vibration, while those at *c* and *g* have reached their lowest points. At points lying between *a* and *c* the particles have various displacements, but all are moving upward; the particles at *d* and *h* are moving downward; and so on. Looking at this figure and thinking of these motions, one can see that in a short time the particles at *a* will have moved downward and those at *b* will have moved up and formed a crest, with the result that the crest will have moved to the right. In the same time *c* has moved up and *d* down, so that the trough has also moved to the right. Thus by a transverse vibratory motion of the cord the crests and troughs travel to the right. *There is no motion of the cord to the right.* It is an important fact in wave-motion that the medium is not carried along with the wave.

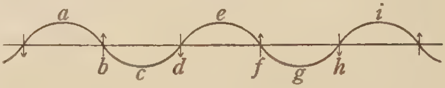


Fig. 181

The particles at *a*, *e*, and *i* are in the same *phase* (sect. 117). Those at *b* and *f* are also in one phase, but opposite in phase to those at *d* and *h*.

A *wave-length* is the least distance between particles which are in the same phase of vibration. The distances *ae*, *bf*, *cg*, are wave-lengths.

The form of apparatus known as Kelvin's wave-apparatus is useful for showing a transverse wave. Rods with heavy balls on their ends are clamped to a steel wire or ribbon (Fig. 182), which is hung from the ceiling. When seen at rest and "edge on," one row of balls lies in a vertical straight line. But when the balls in this row are vibrating with the same amplitude and frequency but with a progressive change of phase, a wave is traveling up or down (Fig. 183). The student should try to visualize this to-and-fro vibration and to see how the crests travel up and down. In what direction must the balls of Fig. 183 be moving in order that the wave may be traveling upward?

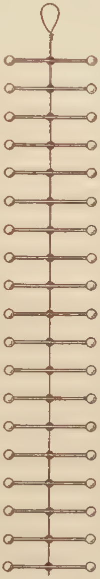


Fig. 182

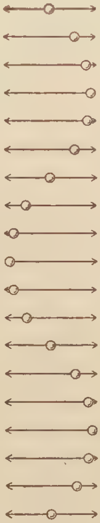


Fig. 183

312. Longitudinal waves. A simple model for longitudinal waves is a long spiral spring hung from the ceiling. Small pieces of white cloth tied at frequent intervals increase the visibility of the motions. If a short portion of this spring is compressed by the hands and then suddenly released, a wave of compression will travel along the spring; or if a part is stretched and suddenly released, a wave of stretch, or extension, will travel up and down the spring. In each of these cases the motion is parallel to the length of the spring. It is therefore a longitudinal motion.

It is usually difficult to follow with the eye the motions of an actual wave. Hence it is best to observe the motions in a slow-moving mechanical model or to study a figure like Fig. 184. The top line of this figure represents an *instantaneous* view of the particles in a longitudinal wave traveling to the right. For convenience only a few particles are shown. Each of these particles is vibrating to and fro in a horizontal direction, the zero position for each being shown by a short vertical line. The arrows show the direction of motion of the particles. At this particular time the first and ninth particles are at a maximum displacement and are momentarily at rest. The third and eleventh particles are in "compressions," while the seventh is in a "rarefaction." From the first to the ninth or from the third to the eleventh is one wave-length. The second horizontal line shows the position of the *same* particles one eighth of the time of a complete vibration (or period) later. As shown in this line, the condensations and rarefactions have moved to the right during this interval of time. The third line shows the particles at a time two eighths of the period later than the top line, the fourth at a time three eighths of the period later, and so on. The last line gives the position of the particles after they have made one complete vibration. It shows that the condensations and rarefactions have moved one complete wave-length to the right.

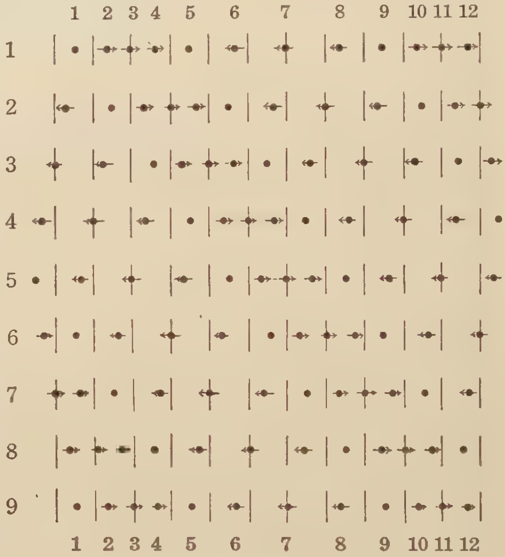


FIG. 184

Careful study of Fig. 184 will show a number of facts regarding longitudinal waves. Among them may be mentioned these:

1. Each particle is vibrating to and fro with the same amplitude and period.

2. The particles in the wave do not vibrate together; there is a progressive change in phase from one particle to the next.

3. Most important is it to see that the to-and-fro motion of the particles produces a movement of the condensation and the rarefaction to the right.

4. In the condensation the particles are swinging to the right, in the direction the wave is traveling.

5. In the rarefaction the particles are moving to the left, opposite to the direction the wave is traveling.

6. If it be imagined that these particles are connected by spiral springs (in a real wave there is something equivalent to an elastic

bond between them), each may be regarded as exerting a force on the one in front of it. In the condensations each spring will be pushing the particles in front forward; but in the rarefactions, where the springs will be stretched, they will be pulling the particles in front back. Since the particles move in the direction of these forces (see 4 and 5 above), each spring is doing *work* on the particle ahead of it. It is an important fact, true in all wave-motions, that there is always a transfer of energy in the direction the wave travels.

313. Transfer of energy in a wave. While there is no transfer of the medium in the direction a wave is traveling, there is always a transfer of energy in the direction of the propagation of the wave. In the preceding paragraph, attention was called to the fact that the forces and motions of the vibrating particles are always in such a direction that each particle is doing work on the one in front of it. The same thing is true of the forces and motions in transverse waves. Referring to Fig. 181, it will be seen that the segment of the cord at *e* is pulling upon *f* in the direction *f* is moving; hence it is doing work on the part at *f*. The segment at *e* also exerts a force on *d*; but *d* is not moving in the direction of the force, and hence does not receive energy from *e*. Energy is transferred only forward.

Wave-motion is one of the important methods of the transfer of energy. For example, it is the only method by which the earth receives energy from the sun.

314. Water-waves. The waves along the surface of the water are a combination of the two types of waves just discussed. Each

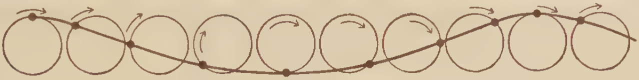


FIG. 185

particle of the water moves up and down (that is, transversely), and at the same time moves to and fro in the direction the wave is traveling. Water is flowing in the direction of the wave-motion on the crest and in the opposite direction in the trough. The resultant motion of any particle (for the case of deep water) is circular. In Fig. 185 are shown the paths of a number of particles which lie on the surface of water, and the wave-form resulting from their motions. Each of the particles is rotating in a clockwise circular motion. One can see from this figure, by imagining the position of the particles a short time later, that both the crest and the trough are moving to the right. Here, as in other cases, while the wave moves to the right, there is no general movement of the water in that direction.

The speed of a surface wave in a deep liquid can be computed from the following formula:

$$V = \sqrt{\frac{g\lambda}{2\,\pi} + \frac{2\,\pi S}{\lambda d}}, \qquad (1)$$

where g is the acceleration of a falling body, λ the wave-length, S the surface tension, and d the density of the liquid. For water-waves longer than about 10 centimeters the first term on the right is so much larger than the second that the last may be neglected entirely; but for ripples which have a wave-length less than about 2 millimeters the first term may be neglected and only the second used. In long waves gravitational force predominates, but in short waves surface tension is the more important factor.

When the depth of water is small compared with the wave-length, the motion of the particles is not in simple circular orbits, as it is in deep water. In the case of shallow water it has been found that for waves longer than about 10 centimeters the speed is independent of the wave-length but varies inversely as the square root of the depth. The formula is

$$\text{Speed in shallow water} = \sqrt{gh}, \qquad (2)$$

where g is the acceleration of gravity and h the depth.

Equation (2) enables us to point out one thing which tends to make sea waves curl over and break as they travel toward a beach. Since each crest is in deeper water than the trough in front of it, it travels faster than the trough and ultimately falls over into the trough.

The formula for short waves in shallow water is

$$\text{Speed of short waves in shallow water} = \sqrt{\frac{2\,\pi h}{\lambda}\left(\frac{g\lambda}{2\,\pi} + \frac{2\,\pi S}{\lambda d}\right)}, \qquad (3)$$

where the symbols have the same meaning as in equations (1) and (2).

The derivation of these equations will be found only in advanced treatises.

315. Relation between velocity, period, frequency, and wave-length. The *period* of vibration has been defined as the time required for a complete vibration. It is also the time required for a wave to travel one wave-length. Why? The latter way of thinking of the period is useful in getting some needed quantitative relations.

Since T, the period, is the time required for a wave to travel one wave-length, a distance λ, the velocity of the wave, is given by

$$V = \frac{\lambda}{T}. \qquad (4)$$

The frequency n (that is, the number of vibrations in 1 second) is equal to 1 divided by the time of 1 vibration, or

$$n = \frac{1}{T}. \qquad (5)$$

Substituting this in equation (4),

$$V = n\lambda, \qquad (6)$$

a very useful equation.

There is a simple way of looking at equation (6). The frequency n is the number of waves given off from a source in 1 second; λ is the length of each of these waves. The product $n\lambda$ is the total distance that the wave will travel out from the source in 1 second. It is therefore equal to the velocity.

316. Standing waves. One can easily make a rope or a rubber tube, fastened at one end, vibrate in segments as shown in Fig. 186.

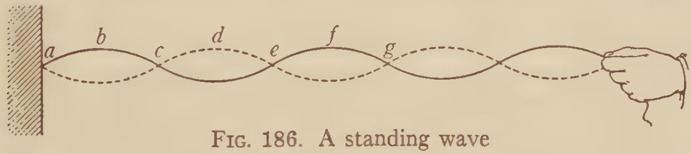


FIG. 186. A standing wave

The parts a, c, e, and g, which remain at rest, are called *nodes*; the parts b, d, and f are called *loops*. This sort of wave is called a *standing*, or *stationary*, wave.

Standing waves are very common; in fact, most of the cases of the vibrations of bodies are instances of standing waves. The swaying of a tall structure or of a bridge, the vibrations of a tuning-fork, the vibrations of the strings of instruments, the vibrations of air columns in wind instruments, the electrical surges in long wires, are all cases of standing waves.

Standing waves are produced when two trains of waves of the same wave-length and amplitude travel through the same medium in *opposite* directions. To produce standing waves in a rope, such as are represented by Fig. 186, the hand is used to send out a train of waves which are reflected back at the farther end, with the result that two trains of waves travel in opposite directions along the rope.

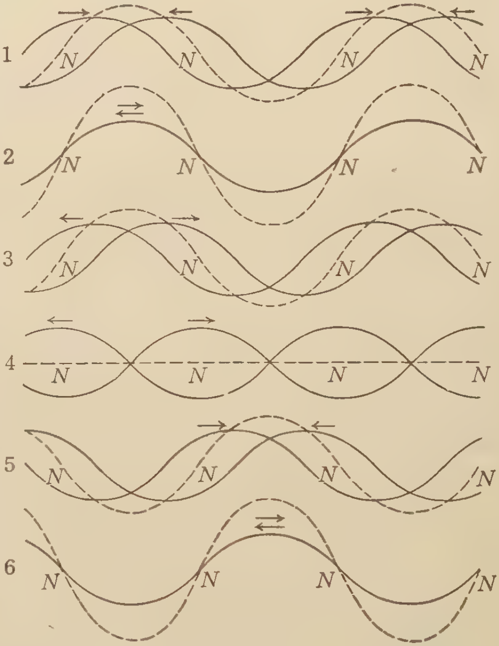


FIG. 187

In Fig. 187 is shown graphically the addition of two wave-trains which are traveling in opposite directions. The broken line is the resulting wave, found by plotting the points which are the sum of the displacements of the two waves. The second group in the figure shows the two waves and their resultant one eighth of a period later than the first; that is, each wave has advanced one eighth of a wave-length. In this case the two waves coincide, and their resultant has twice the amplitude of each. The third group shows the waves one eighth of a period later than the second, and so on, each group showing the waves advanced one eighth of a wave-length farther than the preceding one. The figure thus shows the resulting disturbance for six cases at intervals of one eighth of a period. In the second group the resulting disturbance is greater than in the first, having reached its maximum value. In the third it has become smaller. It is zero in the fourth. In the fifth the displacements are in opposite directions. The maximum displacement in the opposite direction is reached in the last group. If additional figures were given, the vibration could be followed farther. This, however, should be enough to show clearly that the resultant disturbance is a standing wave, with nodes at points marked N.

It should be evident from the figure that *the distance between two adjacent nodes of a standing wave is equal to half a wave-length of the moving waves.*

317. Sound. Sound is sometimes defined as a sensation, but in this text it will be regarded as an external physical disturbance. In order that this disturbance can be propagated two things are essential:

1. There must be a vibrating source.
2. There must be a medium which will transmit, in the form of a wave-motion, the disturbance produced by the source.

It is usually a simple matter to show that there is motion in a source of sound. The vibrations of the strings of musical instruments, of the diaphragm of a telephone receiver, or of the diaphragm of a phonograph are easily detected.

The necessity of a medium to convey the disturbances from a source can be illustrated by putting an electric bell under a bell jar of an air pump. The purpose is to show that no sound can be heard when the air is removed. But air is not the only medium by which the disturbances can be conveyed from the bell. Hence the bell must be suspended inside the jar and not permitted to touch the glass bell jar or the plate of the air pump. When the bell is well isolated, and a good vacuum obtained in the jar, the sound from the bell becomes inaudible.

Air is the usual medium for carrying sound to the ears, but it is well known that sound can be carried through liquids and solids. During the World War, sounds propagated through water were used to detect the positions of submarines. The American Indian was able to hear some sounds through the ground more distinctly than through the air. Distant sounds in a building may be carried to us chiefly by the walls and floors.

318. Sound-waves longitudinal. In order that any medium may transmit waves it must have elasticity; in other words, there must be forces which tend to restore the medium to its original position. When a transverse wave is traveling through a medium, there is a distortion of the medium, a change in the *shape* of any given volume of it. Hence to transmit a transverse wave there must be *elasticity of shape*. In the case of a longitudinal wave there are compressions and rarefactions of the medium; that is, changes in its density. Hence any medium which resists compression and rarefaction elastically can transmit longitudinal waves. The medium must have *volume elasticity* (sect. 178).

Gases and liquids have only one kind of elasticity—volume elasticity; therefore gases and liquids can transmit *only longitudinal or compressional waves* (an exception must be made in the case of waves on the surface of liquids, where the forces of gravitation and surface tension resist changes in the shape of the surface).

The air in a sound-wave vibrates to and fro longitudinally as in Fig. 184, sect. 312. But sound-waves do not travel along one line of particles only: they usually travel out in all directions from their source. Hence the source is surrounded by moving spherical layers, alternate ones being condensations, the others rarefactions. In accordance with 4 and 5, sect. 312, in the condensations the air is moving away from the source; in the rarefactions, back toward the source.

In the case of very intense waves the changes in the density of the air may be considerable. Because of the increased density, light passing through a condensation may be bent, or refracted (Chapter XLII). For this reason intense condensations may cast rather sharp shadows. These have been photographed under special circumstances. In several instances in France, during the World War, long, narrow "shadows" of sound waves were seen moving across the clouds in a direction perpendicular to their length.

319. Speed of sound. It is a matter of common knowledge that it requires time for a sound to reach our ears. The steam from the whistle of a distant engine can be seen before the sound is heard. A baseball can be seen to leave the bat before we hear the sound of the blow. In the case of the 220-yard dash the timers, who stand at the finish, do not wait until they hear the sound of the starter's gun, but start their watches when they see the smoke.

Thunder is the sound produced by a flash of lightning. But the sound is always heard after (often several seconds after) the flash. (The danger is always over when we hear the thunder.)

A simple method of determining the speed of sound is to measure by a stop watch the interval between the time when the flash of a gun at a known distance is seen and the time when the sound is heard. In accurate work allowance must be made for wind, temperature, and humidity in the air. Great pains have been taken to determine accurately the speed in air. It is now quite certain that the exact value of the speed in dry air at 0° C. is between 331 and 332 meters per second.

SPEED OF SOUND

(In Meters per Second)

Air at 0° C.	331	Sea water at − 0.3° C.	1453
Hydrogen at 0° C.	1258	Steel (soft)	4700–5200
Oxygen at 0° C.	317	Copper	4000
Water at 13° C.	1441	Glass (soda)	5000–5300
Iron (wrought)	4900–5100	Glass (flint)	4000 (approximately)

It has been proved very accurately that the speed of compressional waves in air is independent of their wave-length. In musical terms this means that the speed is independent of the pitch. If there were much difference in the speeds, the high-pitched notes of a distant band would not reach the ear at the same time as the low-pitched ones, with the result that the sounds would not be harmonious.

A very remarkable compressional wave was produced by the explosion of the volcano Krakatoa, in the East Indies, in 1883. This wave went around the earth, converging at a point opposite Krakatoa. There it was reflected back to its source, then back again, being detected seven times at some places. The wave was so long and the change in pressure so slow that it affected recording barometers. From the barometric records the velocity was found to be about 320 meters per second. This agrees fairly well with the velocity of sound-waves, when the inaccuracies of the method and air temperatures are taken into account.

The speed of sound-waves in air is independent of changes in barometric pressure, but it does depend on the temperature of the air. The speed of sound increases approximately 60 centimeters per second for each degree centigrade of rise in temperature. The speed in dry air is given more accurately by the equation

$$V = V_0\sqrt{1 + 0.00367\,t}, \qquad (7)$$

where V is the speed at the temperature t degrees centigrade and V_0 the speed at 0° C.

In general, the speed of compressional waves in any medium is given by the equation

$$V = \sqrt{\frac{E}{d}}, \qquad (8)$$

where E is the modulus of volume elasticity and d the density. The reason that the speed is so great in liquids and solids is that E is so large for those substances. The quantity E has the same value for the gases hydrogen, nitrogen, and oxygen when they are at the same pressure; hence in these gases the velocity of compressional (sound) waves varies inversely as the square root of their densities.

320. Determination of the direction of a source of sound. All of us learn early in life to locate the direction of a source of sound, but just how we do it is not definitely known. Apparently the process is a complicated one, with a number of different factors taking part. It is known that at least two different methods are used,—as a rule, simultaneously.

1. The *difference in the intensity* of the sound at the two ears is known to help us in locating a source of sound, especially if the head is moved. Several things tend to produce this difference in the intensity at the two ears when the sound is not directly in front or behind. The chief cause is the "sound-shadow" produced by the head. This method of locating the source is more effective for short waves—those from sounds of high pitch. Since long waves bend around obstacles more readily than short waves, there is less difference in the intensity of long waves at the two ears.

2. When the two ears are at unequal distances from the source of sound, condensations and rarefactions reach one ear before they do the other, with the result that the vibrations in the two ears are not "in step"; that is, they are not in the same *phase*. This difference of phase of the vibrations at the two ears is an important factor in the location of sounds. For sounds which are no higher in pitch than the human voice (that is, for waves longer than about 27 centimeters) this factor is more important than the first one, especially when the head is not moved.

The principle of the second method was used in the World War for detecting the direction of submarines and invisible airplanes. In the method used for the location of airplanes two sound-receivers A and B are mounted at the ends of two tubes, as shown in a simplified form in Fig. 188. The other ends of the tubes are connected to the ears by tips similar to those used in a physician's stethoscope. The "horn" A is connected to one ear, and B to the other. If B is nearer to the source of sound than A, the waves entering at B and reaching one ear are ahead in phase of those reaching the other ear by way of A. If in this case B is connected to the right ear, the sound will appear to come from the right. If the apparatus is rotated until the sound seems to the listener to be directly in front of him, A and B will be equally distant from the source. In this way the direction of the source can be accurately determined.

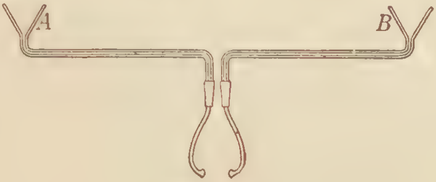


FIG. 188

Apparatus similar in principle was designed for the location of underwater sounds, such as those produced by a submarine. Instead of the "horns" A and B, the submerged ends of the tubes were covered with rubber nipples, which transmitted the sound-waves from the water into the air in the tubes. But because it was not convenient to rotate a ship to which they were attached, and for other reasons, a more complicated apparatus was developed. However, the principle of locating the direction by the difference in phase of the waves transmitted to the two ears of the listener was the basis of the more involved apparatus. A high degree of accuracy was attained by these instruments.

321. Sound-ranging. During the World War a number of methods were devised for locating, by means of sound-waves, the position of large guns. One method is known as the method of the tangent circle (Fig. 189). The time when the sound reaches each of the stations A, B, and C is accurately observed. Let the time at A be called t_1; that at B, t_2; and that at C, t_3. The arcs of two circles are drawn on a map or suitably prepared chart. The radius of the circle used at B is computed from the equation

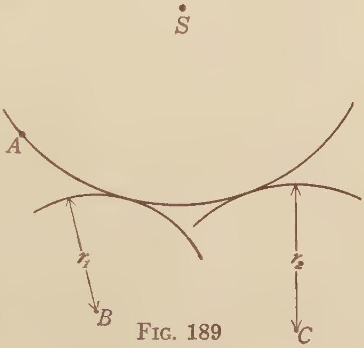


FIG. 189

$$r_1 = V(t_2 - t_1),$$

where V is the speed of sound. The radius used for the circle about C is computed from

$$r_2 = V(t_3 - t_1).$$

Then a circle drawn through A, and tangent to the arcs about B and C, must have as its center the position of the source of the sound-wave. That this is true is clear when the physical meaning of r_1 and r_2 is realized. The first of these, r_1, is the distance sound travels in the time $t_2 - t_1$. It is therefore the difference between the distances of B and A from the source. Similarly, r_2 is the amount that C is farther from the source than A. The accuracy with which a source of sound several miles away was located by this method was remarkable.

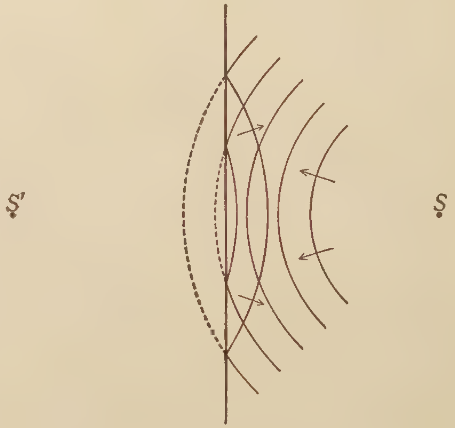


Fig. 190

322. Reflection of sound-waves. The reflection of waves from a source S by a plane surface is shown in Fig. 190. Since reflection does not change the velocity of the waves, a reflected wave is at any instant just as far in front of the plane as it would have been behind if the plane had not interfered. The figure shows that after reflection the waves diverge from a point, S', which is as far behind the plane as the source, S, is in front. Many different kinds of waves are reflected in this manner—water-waves, light-waves, and sound-waves.

Many commonly observed phenomena are explained by the reflection of sound-waves. A curved surface will reflect sound-waves in a manner similar to the way a curved mirror reflects light. If a watch is hung at the focus of a large concave mirror, the ticking may be heard many feet away. Rooms or halls with arched ceilings produce the well-known "whispering gallery" effect.

Echoes are well-known instances of reflection of sound. They may be produced by buildings, masses of trees, cliffs, or anything which serves as a sound-reflector. When more than one thing reflects the sound, multiple echoes may be heard. A sound of a gunshot in a canyon or on a lake surrounded by cliffs may be drawn out into a long roar by the reflections from different surfaces. The long, rolling sound of thunder is due partly to reflections from different cloud masses and layers of warmer or colder air.

Recently there has been developed a rapid method of deep-sea sounding in which reflected sound-waves are used. The waves travel downward from a ship and are reflected from the bottom of the sea. The interval of time between the starting of the wave and the instant when it returns is measured. From the known velocity of propagation and this time interval the distance to the bottom is computed. During the World War sensitive instruments were developed for the detection of sound-waves in water. These have been further improved, and this deep-sea sounding is only one of their many uses.

In a room most of the sound we hear has been reflected. In a large auditorium the reflection of sound, first by one wall and then by another, produces a reverberation which when long continued makes it difficult to understand the speaker or to get the harmony of music. There should be some reverberation; otherwise the effect would be the same as out of doors, where the lack of reflection makes it harder to hear a speaker. The proper amount of reverberation—not too little nor too much—is necessary to make the acoustic properties of an auditorium correct. Too often the subject of architectural acoustics is neglected in designing public halls and auditoriums. Fortunately, much work has been done along these lines, and there are now experts who, from the plans of an auditorium, can predict accurately its acoustic properties.

323. Refraction of sound-waves. Those who have bathed on ocean or lake beaches may have observed that no matter in what direction waves are traveling in deep water, they always roll up nearly parallel to the beach. Photographs of scenes on a beach usually show the waves parallel to the shore line. The waves coming in from deep water are always bent, or deflected, as shown in Fig. 191. The cause is that the speed of waves in shallow water varies with the depth: the more shallow the water the slower is the motion (sect. 314). As seen in the figure, part of each wave is in deeper water, where it travels faster. The result is that the wave swings around as it approaches the shore. Since these waves always travel in a direction perpendicular to themselves, the direction in which they travel is changed as shown by the arrows in the figure. Changes in the direction of wave-motion are often produced by changes in velocity. Such a change in direction is called *refraction.*

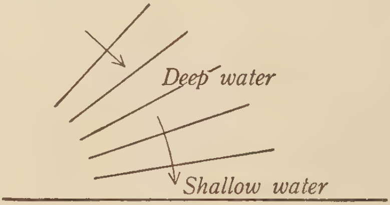


Fig. 191

Refraction of sound-waves may be produced in several different ways, one of which is the very common effect produced by wind. Usually the wind has less velocity near the ground than it has higher up. Since the wind changes the effective speed of sound, a sound traveling against the wind is refracted upward (Fig. 192), and sound traveling with the wind is refracted downward (Fig. 193). In the last case the sound is refracted down to the ground, reflected up, then refracted down again, and so on. The result of this is that sound travels much farther with the wind than against it. Some remarkable instances of sound carried over a mountain and down into the next valley can be explained by this refraction.

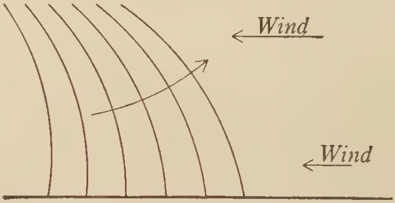


Fig. 192

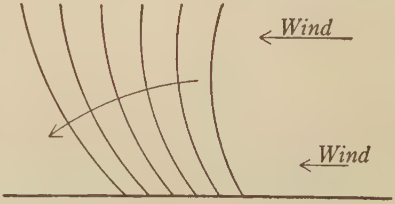


Fig. 193

Sound travels faster in warm air than in colder air. In the daytime the air is usually warmer near the ground. The result is that the speed of sound is greater near the ground. A sound-wave will then be refracted upward, as in the case of Fig. 192. At night the air is often cooler near the ground than it is at greater elevations. Sound-waves will have a smaller speed near the ground, and the waves will be refracted as in the case of Fig. 193. We should therefore expect—what is a well-known fact—that sound-waves would travel farther along the ground at night than in the daytime.

The effect of reflection and refraction of sound-waves is a matter of great importance in the case of fog signals. Rising columns of warm air or descending currents of cold air not only refract these signals but also reflect them. Sometimes these effects reduce the intensity of the sound so much that the signals are heard only at relatively short distances.

During the World War there were many strange things connected with the sounds from the heavy guns. At times the sound of the firing would be heard at places a hundred and fifty miles away and yet be entirely inaudible in a wide zone nearer the source. Many of these effects have not been completely explained, but they are undoubtedly instances of refraction or reflection, probably both.

324. Musical sounds. Musical sounds are distinguished from mere noises by the fact that a musical sound has a definite pitch. The distinction is not always easy to make. A certain sound may at one time seem to be a noise, and at another time its pitch may be easily recognized. One can cut out a number of pieces of wood, about five sixteenths by three quarters of an inch in section and of lengths varying from about eight inches to five and a half, which, when dropped on a table, one after another, will give a musical scale. But if one hears only one stick, the sound would probably be called a noise. In noises the pitch, if there is any, is usually masked.

Musical sounds have three characteristics: *pitch*; *intensity*, or *loudness*; and *quality*, or *timbre*.

325. Pitch. By the aid of a well-known experiment a simple fact regarding pitch can be learned. A toothed wheel is mounted so that it can be rotated at different speeds. A card or a piece of paper, held so that each tooth hits it, is set in vibration by the rotating wheel and gives off a sound of definite pitch. When the wheel is speeded up, the rate of vibration of the card is increased, and, judged by the ear, the pitch of the sound is raised; but when the speed of the wheel is decreased, the pitch falls. This shows in a very simple way that *pitch depends on frequency of vibration:*

the higher the frequency, the higher the pitch; the lower the frequency, the lower the pitch.

The change in pitch effected by a change in frequency can be observed in many ways. The pitch of the note given out by an electric motor, an electric fan, or a circular saw rises as the speed increases, and falls as the speed decreases. It is well known that in order to raise the pitch of the sound given off by a phonograph the speed of the machine must be increased.

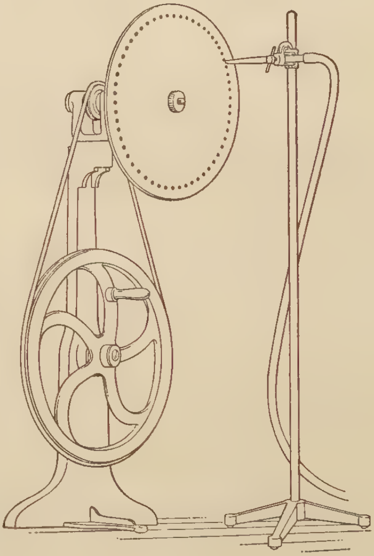


Fig. 194

The siren is an instrument used either to measure the frequency of a musical sound or to produce one of known frequency. The simplest form is a disk containing in a circular row a large number of holes (Fig. 194). This disk is rotated in front of a nozzle from which a jet of air is flowing. The jet is placed so that the air can flow through the holes as they pass across the jet. Each time a hole passes into the jet a puff of air goes through the disk. The frequency of the sound produced by these puffs depends on the number of holes in the disk and on the speed at which the disk is rotated. To determine the frequency of a given sound the speed of the siren is changed until it gives the same pitch as that of the given sound. The number of revolutions per second of the siren disk is then measured. The product of this number and the number of holes gives the desired frequency.

Since the pitch of a note depends only on the frequency of vibration, it is customary to measure pitch by the frequency. The piano has a range of pitch from 27.2 to 4138.4 vibrations per second. A low bass voice can produce a note of pitch 60, and a high soprano voice a pitch of about 1300.

By equation (6), sect. 315,

$$V = n\lambda,$$

where V is the speed, n the frequency, and λ the wave-length. Since the speed V in air is the same for all sound-waves, it follows from this equation that in air the wave-length varies inversely as the frequency. Sounds of high pitch have short waves, and those of low pitch have long waves.

326. Intensity, or loudness. It has been demonstrated that within rather wide limits the loudness of the sound-sensation in the ear corresponds closely to the energy in the wave. For this reason the intensity, or loudness, of a sound-wave is measured in terms of the energy of the wave.* The amount of energy in sound-waves is usually rather small. One joule per second—that is, 1 watt—converted into an ordinary-pitched sound can be heard a mile.

327. Tone quality. Two musical sounds may have the same intensity and the same pitch; yet they will not produce the same effect in the ear if one of the sounds is from a violin and the other from a piano. The characteristic which enables one to tell the difference between these two notes is called the *tone color*, the *tone quality*, or, more simply, the *quality*. The French word *timbre* is frequently used.

Practically all musical sounds consist of mixtures of notes of different pitches. A vibrating string gives out, in addition to what is called the *fundamental*, a number of higher-pitched sounds which are called *overtones*. The tone quality is determined by the number and relative intensities of the overtones. For example, a string on a certain instrument may give, in addition to the fundamental of 200 vibrations per second, an intense overtone of 400 vibrations per second; one of 800 vibrations per second, with a moderate intensity; another of 1200 vibrations per second, with less; and, very faintly, others of 1600 and 2000 vibrations per second. A string on a different instrument may give the same fundamental, and the notes of 400 and 800 vibrations rather faintly, but make the higher-pitched overtones relatively loud. These two tones will differ in tone quality. Whenever the qualities of two tones are different, the overtones in the two are different either in pitch or in loudness, or in both.

---

*Other things being equal, the energy is proportional to the square of the amplitude of vibration; hence the intensity or loudness of a wave depends on the square of the amplitude.

A tuning-fork mounted on a "resonating box" usually gives a pure tone, only the fundamental being present. A pure tone is dull and colorless. The richness of the tones of musical instruments and of the human voice is due to the overtones. Generally speaking, the more overtones present, the richer the tone quality.

When the frequencies of the fundamental and overtones are proportional to the numbers 1, 2, 3, 4, 5, etc., they form a *harmonic series*. The fundamental and overtones in this case are called *harmonics*. For example, when the frequencies of the fundamental and overtones are given by the numbers 220, 440, 660, 880, 1100, etc., they form a harmonic series. In most musical instruments the fundamental and overtones are harmonics; but in some cases they are not. The clang tone of a tuning-fork which has been struck with a block of wood or with some other hard substance is due to an overtone which is not a harmonic.

The intensity of one or more of the overtones is sometimes greater than the intensity of the fundamental. This is true of the notes produced by the *G* string of a violin and of the vowel sounds given by the voice.

The difference between the different vowel sounds as intoned by the voice is entirely a difference in the overtones present. Independently of the pitch at which the vowel *a* in the word *father* is pronounced, it is characterized by an overtone which is much more intense than the fundamental and has a pitch of about 925 vibrations per second. The vowel sound in *gloom* has a characteristic overtone with a pitch a little higher than 300.* Some vowel sounds have two intense overtones; for example, that in *bee* has one pitch of about 300 vibrations per second and another of about 3000. Thus the chief difference between the vowel sounds in *bee* and in *gloom* is in a high overtone with a pitch of about 3000. High-pitched notes have trouble in getting through in a telephone line. Hence, as these two vowel sounds are heard on a telephone, there is little difference between them. The reason that a telephone operator trills the *r* in *three* is to enable one to distinguish that number from *two*. Another interesting thing, which is readily explained by facts just given, is that slowing up a phonograph changes the vowel sounds. For example, the vowel sound in *father* is due to an overtone of pitch of about 925 vibrations per second, that in *raw* to one of about 735, that in *no* to one of about 460, and that in *gloom* to one of about 300. Hence, on slowing down the speed of a phonograph the vowel sound in *father* will be changed to that in *raw*, then to that in *no*, and, at a still slower speed, to that in *gloom*.

---

*The methods used in determining these facts and many other interesting ones are given in D. C. Miller's "Science of Musical Sounds."

328. Resonance, or sympathetic vibration. A child soon learns that in order to "work up" in a swing, he must time the forces he exerts to the natural period of vibration of the swing. Whenever the impressed forces are applied in unison with the natural vibration of a body, larger vibrations are produced. These are called *sympathetic vibrations*, or there is said to be *resonance*. If two balls (pendulums) are hung by strings of the same length from a support which is not rigid, and if one of the balls is given a vibratory motion, the other will be set in vibration by the small disturbances propagated through the support. A convenient support is a horizontal meter stick, narrow edge up, clamped at one end to a table top. The two pendulums are hung near the free end of the stick, 15 or 20 centimeters apart.

Often, when one speaks or sings near a piano, a string may be set in audible vibration. This is due to the fact that the frequency of the sound-waves striking the string is equal to the natural frequency of the string. Although the air vibrations in the sound-wave exert very small forces on the wire, yet because the forces are exerted several hundred times in unison with the movement of the string, a large amplitude of vibration is produced in the string. A rather striking experiment is one in which a tuning-fork resonates to a sound produced by another tuning-fork, of the same pitch, twenty or thirty feet away. If the sound produced by the resonating fork is not audible throughout the room, the vibrations can be made "visible" by a varnished pith ball suspended so that it just touches one prong of the fork.

The sounds given out by most wind instruments are from air columns vibrating in their natural frequency. Because they have natural periods of vibration, resonance can be shown by such a column. If a glass or metal tube is fitted with some sort of piston so that the length of the air column in the tube can be changed, it is easy to find a length for which strong resonance is produced when a tuning-fork is held at the open end. The volume of air in the box on which a tuning-fork is mounted is always made of such a size that there is resonance to the fundamental tone of the fork. The air in an organ-pipe will resonate to a tuning-fork which has the pitch either of the fundamental or of any of the overtones which are emitted when the pipe is sounded. This is one way of showing that there are different modes in which an air column can vibrate. Many other cases of resonance, not only in sound vibrations but in slow, visible vibrations, such as those of bridges, automobiles, and street cars, are frequently observed.

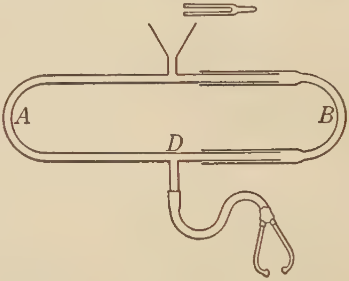


Fig. 195

329. Interference. The apparatus shown in Fig. 195 shows an important property of wave-motion which is called *interference*. The sound-waves from a tuning-fork enter a tube, pass by two different routes A and B, unite, and finally reach the ears through tips such as are used in a physician's stethoscope. The length of the path B can be altered, for the straight tubes on that side slide inside one another. If one listens to the sound transmitted through the tubes while the side B is slowly lengthened, there will be times when no sound is heard,—when the waves traveling through B interfere with those through A. This occurs when the path around B is longer than that through A by an amount equal to half a wave-length, or to three halves, five halves, etc. of a wave-length. For these distances condensations coming through A meet rarefactions coming through B, and rarefactions coming through A meet condensations coming through B. When this occurs the resulting disturbance at the junction point D will be a minimum. The waves are said to *interfere* destructively with each other.

The apparatus of Fig. 195 can be used to measure the wave-length of a sound. When the length of the side B has been changed from a place corresponding to a minimum of sound in the ears to the next place of minimum, the change in length of the path of the sound through B has been 1 wave-length. Since the change of length of B can be measured, the wave-length is directly obtained.

The two prongs of a tuning-fork move in opposite directions. While they are moving toward each other a condensation is produced in the region between them (at B in Fig. 196, where only the ends of the prongs are shown). At the same instant, rarefactions are produced at A and C. On the return swing, while the prongs are moving apart, condensations are produced at A and C, and a rarefaction at B. The waves from B and those from A and C thus start in opposite phases and interfere with each other along the dotted lines in the figure. These regions of silence are easily detected if a vibrating fork is held near the ear and slowly rotated.

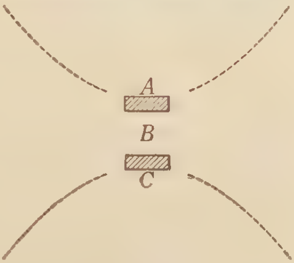


FIG. 196

Interference of waves has many important applications in the subject of light. This is much more fully discussed in Chapter XLVI. Another example of the interference of sound-waves is given in the following section.

330. Beats. When two men are walking together, one taking a slightly shorter step than the other, they will be at one time in step, later out of step, then in again, and so on. If in one minute one man takes exactly 70 steps and the other 68 steps, the first will gain 2 steps on the other every minute. They will thus be in step at regular intervals, the length of each interval being the time required for the first man to gain a whole step. Since he gains 2 steps each minute, they will be in step twice each minute. From this the following rule may be formulated: *If two men with different gaits walk together, the number of times they are in step in unit time is equal to the numerical difference of the frequencies of the steps.*

We may apply the foregoing principle to sound-waves emitted by two tuning-forks of different frequencies. If one fork vibrates with a frequency of 256 vibrations per second, and the other with a frequency of 260, the forks will be vibrating together four times each second. Four times during each second the condensations from the two forks will reach the ear simultaneously, and four times each second the condensations from one fork will reach the ear at the same time as the rarefactions from the other fork. When opposite in phase, the two waves *interfere* with each other; when in the same phase, they *reënforce* each other. This produces a phenomenon known as *beats*, a periodic change in the intensity of the sound. Four times each second the sound is a maximum and four times it is a minimum; or, as usually stated, there are four beats heard each second. From the reasoning given in the early part of this section a general law of beats may be stated, as follows: *The number of beats heard each second is numerically equal to the difference in frequencies of the two sounds.*

Beats are used a great deal in tuning two sources of sound to the same pitch. Two forks can be brought very accurately to the same pitch by adjusting one until each of the two forks produces the same number of beats per second when sounded with a third fork which has a slightly different pitch.

331. Doppler's principle. As a rapidly moving locomotive passes an observer he can detect a change in the pitch of its whistle. If a high-pitched tuning-fork or whistle is rapidly swung toward and away from an observer, changes of pitch can be detected. The change is not very great, however. If an engine is moving at fifty miles an hour, the pitch of the whistle as it approaches is a little more than a whole-tone interval* higher than that heard as it moves away. Or, if one riding in a train notes the pitch of a gong at a grade crossing, he will observe an apparent change in pitch as he passes the gong.

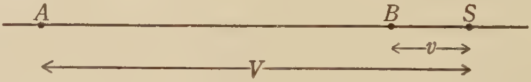


FIG. 197

The reason that the pitch of a note is changed by the motion of the source can be understood by the aid of Fig. 197. Let S be the source, AS a distance equal to the velocity of sound, and BS equal to the velocity of the source of sound. If the source is stationary, the sound-waves given out during a second will, at the end of that second, be spread out over the distance AS; but if during this second the source has moved to B, the waves will be crowded into the distance AB. Hence the effect of the motion is to shorten the emitted wave-length. Since the waves are shortened, the number which pass the ear beyond A each second is increased; for the general relation between the frequency, wave-length, and velocity, given by equation (6), applies here. If n' is the altered frequency and λ' the shortened wave-length,

---

* A whole-tone interval is the interval between the notes C and D, or A and B, on a piano.

$$n' = \frac{V}{\lambda'}. \tag{9}$$

The value of λ' will now be found. Since the number of the waves crowded into the distance AB is the number, n, emitted each second, the length, λ', will be

$$\lambda' = \frac{AB}{n} = \frac{V - v}{n},$$

where V is the velocity of sound and v the velocity of the source. Substituting the value of λ' in equation (9),

$$n' = n\frac{V}{V - v}. \tag{10}$$

This is the formula for computing the frequency of the sound heard when the source is moving toward the observer with a velocity v.

When the source is moving away from the observer, it can be shown by similar reasoning that

$$n' = n\frac{V}{V + v}. \tag{11}$$

In this case the pitch is lowered by the motion.

The pitch is apparently changed also in the case where the source of sound is stationary and the observer is in motion. If the observer is moving toward the source with a velocity v, the pitch is raised. It can be proved that the apparent frequency n' is given by the equation

$$n' = n\frac{V + v}{V}. \tag{12}$$

If the observer is moving away from the source, the pitch is lowered, and the formula for the observed frequency is

$$n' = n\frac{V - v}{V}. \tag{13}$$

332. Vibrating strings. The vibrations of strings of musical instruments are instances of the standing waves of section 316. It is not difficult to show that a string may vibrate in any one of the modes shown in Fig. 198. For example, if a number of light paper "riders" are placed on a string, and if the string is bowed or plucked while a finger lightly touches it at a point a third of the length of the string from one end, the riders will show by their motion where the nodes and loops are. In this case the string vibrates in three equal segments, with one of the nodes at the point where the finger is touching the string. Other modes of vibration may be shown by touching the string lightly at the midpoint or at a point one fourth or one fifth of the length from one end.

When the string vibrates in two segments, it gives a note which has a frequency twice that of the fundamental; when vibrating in three segments, one three times as great; when vibrating in four segments, four times; and so on. As stated in section 327, a string may have all these different modes of vibration simultaneously, and give out a complex tone which is a mixture of notes having frequencies corresponding to those of the string vibrating in the different modes. If the frequency of the string vibrating in one segment is 200 vibrations per second, the tone heard is composed of frequencies 200, 400, 600, 800, etc. The ear can often detect some of these harmonic overtones. When a string is plucked near one end, the higher-pitched notes become very prominent and are easily detected. There are a number of experiments which prove that a string may vibrate in such a complex way. Two of them will be described.

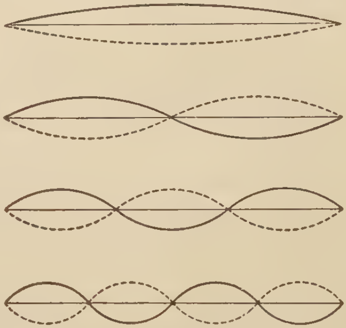


FIG. 198

If a string which has been plucked is lightly touched at its middle point, the vibration of the string as a whole is suppressed, but those vibrations which have nodes at the middle point will continue. The ear hears the pitch of the string change, from that of the fundamental, to a fainter note having twice the frequency of the fundamental. This vibration must have been present before the string was touched. If after it has been plucked, the string is touched at a point a third of its length from one end, the ear hears the pitch changed to a frequency three times that of the fundamental. With a little patience this experiment can be continued, and the existence of other overtones shown.

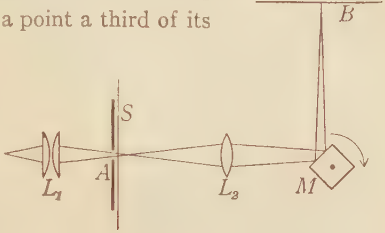


FIG. 199

A more elaborate experiment not only brings out the fact that the vibrations of a string are complex but gives an important way of looking at the subject of complex vibrations. Fig. 199 shows the horizontal plan of the apparatus. In front of the condenser lenses of a projection lantern is placed a vertical slit *A*, and close in front of this a stretched horizontal brass or steel wire *S*. *M* is a cube with mirrors on the four vertical sides, so mounted that it can be rotated about a vertical axis. The lens L_2 is used to form an image of the slit on the screen *B*. When the mirror is stationary, the image of the narrow slit shows a short horizontal black line across it (Fig. 200). The short black line is the image of that part of the wire which lies in front of the slit. When the horizontal wire is vibrating in a vertical plane, the image of the wire will be widened out into a band as shown in Fig. 201. If the mirror *M* is now rotated about a vertical axis, the image of the slit will be moved sidewise across the screen, and the image of the vibrating wire will trace out a wavy line which can be plainly seen. If the wire is vibrating with only a pure tone (that is, in only one mode), the shadow of the wire traces out a smooth sinusoidal curve. Such a vibration was called in Chapter X simple harmonic motion. What kind of curve should we expect if the vibrations of the string are complex; that is, if vibrations in segments are superposed on the fundamental mode? To answer this let us suppose that the piece of wire in front

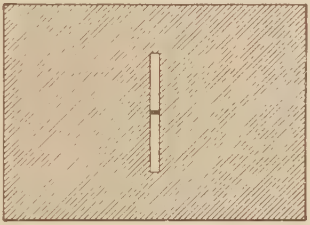


FIG. 200

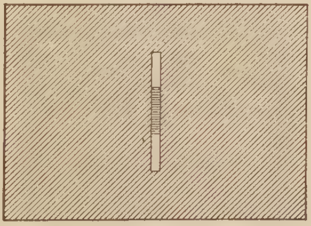


FIG. 201

of the slit is vibrating with a certain frequency, and at the same time with two other frequencies, twice and three times the first. If each component of the vibration could be projected on the screen, one at a time, the curves *a*, *b*, and *c*, in Fig. 202, would be seen. If the string actually has all these vibrations simultaneously, the displacement of any short piece at any instant is the sum of the displacements of the three components. Curve *d*, in Fig. 202, has been drawn so that it is the sum of the three simple curves. Since *d* is the sum of *a*, *b*, and *c*, it is the form of curve we should expect traced out on the screen if the wire has all three vibrations simultaneously.

From the foregoing it appears that if the curve traced out on the screen is a smooth sinusoidal curve, the string is vibrating in a simple mode; but if the curve is complex, as in *d* of Fig. 202, the string is vibrating with more than one frequency. The actual curves obtained from the string are complex, of various forms, depending on how the string is plucked or bowed. Hence this experiment shows us conclusively that a string does vibrate in a complicated manner,— that it has simultaneously vibrations of different frequencies.

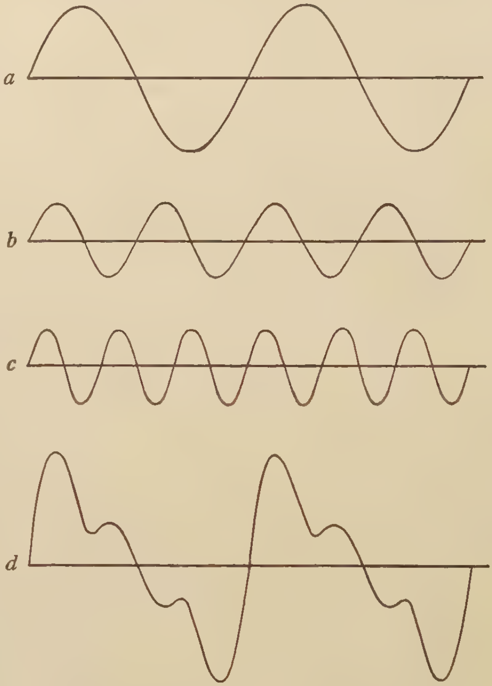


Fig. 202

The following experiment is one of several that can be shown with this vibrating string. If the string is bowed or plucked at different places or in different ways, the ear will detect changes in tone quality. At the same time that the tone quality changes, the form of the curve will change. This shows that the quality is changed when the overtones are changed.

The shape of one of these curves, which show graphically how a *piece* of the string vibrates, is sometimes called the *wave-form*. Such curves not only show how the string must vibrate but also give one an idea of the vibrations of the air in the sound-wave which transmits these vibrations to the ear. The air, however, vibrates longitudinally, not transversely.

333. Frequency of vibration of strings. It is proved in advanced texts that the frequency of the fundamental mode of vibration of a stretched string is given by the equation

$$\text{Frequency} = \frac{1}{2L}\sqrt{\frac{F}{m}}, \tag{14}$$

where L is the length of the string, F the stretching force, or tension, and m the mass per unit of length of the string. If L is measured in centimeters, F in dynes, and m in grams per centimeter length of the string, the frequency will be the number of vibrations per second. It is not a very difficult experiment in the laboratory to determine experimentally the four quantities in equation (14) for a given string, and thus to verify the equation. This is often given as a laboratory exercise.

Equation (14), which is often called the *law of stretched strings*, shows that the frequency varies (1) inversely as the length; (2) as the square root of the tension; (3) inversely as the square root of the mass per unit length of the string.

By adjusting the length or tension, or by using a lighter or heavier string, notes of any desired frequency can be obtained from a vibrating string. For example, in the case of the violin the different strings have different masses: the G string, which gives the lowest frequency, is the heaviest. The strings of this instrument are tuned by changing the tension. The performer can obtain any frequency, within wide limits, by changing with his fingers the effective length of the strings.

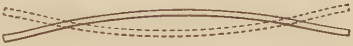


FIG. 203

334. Vibrating rods. In some musical instruments the tones are given out by bars or rods vibrating transversely. When struck near the middle, they vibrate as shown in Fig. 203, with nodes near the ends. They are usually supported at these nodes.

A tuning-fork is a case of a bent bar. It vibrates with two nodes near the bend (Fig. 204). When the prongs vibrate to and fro, the handle moves up and down with a very small amplitude.

A straight rod which is clamped at the middle can be set into *longitudinal* vibrations by stroking it with a piece of leather on which a little rosin has been placed. When the rod is vibrating longitudinally, the two ends of the rod move toward and then away from each other. The middle part of the rod is alternately compressed and stretched. That the end of the rod is really vibrating in a direction parallel to the length of the rod is usually shown in the classroom by suspending a light, varnished pith ball so that it just touches the end of the rod. The longitudinal vibrations of the rod cause the ball to be driven off.

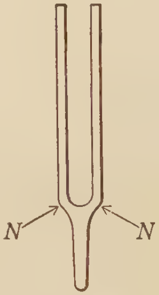


FIG. 204

335. The Kundt's-tube experiment. The velocity of compressional waves in metals can be obtained by an experiment due to Kundt. A rod, R, of the metal to be tested is clamped at the center, with one end extending into a long glass tube, T (Fig. 205). The other end of the tube is closed with a piston, P, which can be moved in or out. On that end of the rod which projects into the tube is fastened a disk of cardboard which nearly closes the tube. When the rod is set into longitudinal vibration by stroking the free end with a piece of rosined leather, longitudinal waves are produced in the air in the tube. These waves, traveling down the tube, are reflected back by the piston P. This gives the necessary condition for standing waves — two wave-trains traveling in opposite directions. If the piston P is moved in or out, an optimum position is obtained where the tube is in resonance and the standing waves are very strongly marked. The effect becomes visible when some lycopodium powder or fine cork filings are placed in the tube. The violent oscillations of the air move the light particles in such a manner that it is easy to locate the nodes and loops of the standing air-waves. The length of the air-wave is twice the distance between adjacent nodes.

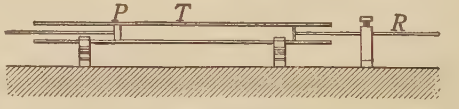


FIG. 205

Since the rod is vibrating with a loop at each end and a node in the middle, the wave-length in the rod is twice the length, L, of the rod. Hence equation (6) may be written

$$V_R = n(2\,L),$$

where V_R is the velocity of a compressional wave in the rod. In the air column

$$V_a = n\lambda_a,$$

where V_a is the velocity, and λ_a the wave-length, in the air. Dividing the first of these equations by the second,

$$\frac{V_R}{V} = \frac{2\,L}{\lambda_a}. \qquad (15)$$

Since L and λ_r are measured, and V_a, the velocity in the air, is known, the velocity V_R can be computed.

Sometimes the experiment is varied by making the glass tube air-tight and putting different gases in the tube. This affords a method for obtaining the wave-length and hence the velocity of sound in different gases.

336. Vibrations of air columns. Vibrating air columns are used so much in musical instruments that the character of vibrations of the air in these cases is of much interest. However, only one case will be considered here — that of long, narrow tubes. It is only in this type that the case is simple enough for our consideration.

The type of vibration in all air columns is that of a standing wave. When it is possible to locate the nodes and loops of the standing wave, the character of the vibration and the wave-length are at once known. When the wave-length is known, the frequency of vibration can be computed by the aid of equation (6). Hence the location of the nodes and loops is the first important step.

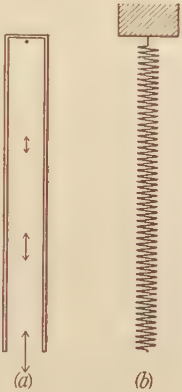


FIG. 206

If the end of a column of air is closed, there must be *a node at the closed end*, since the air next to this end cannot move. At the open end of a tube the condition is quite different. Here the air has greater freedom of vibration than at any point within the tube. At the *open end* is a loop. In Fig. 206, (a), is shown diagrammatically a vibrating column of air with a node at the closed end and a loop at the open end. The arrowheaded lines show by their length the relative amplitudes of vibration of the air. A suspended spring (see Fig. 206, (b)) will vibrate up and down with the same sort of motion as the air in the tube,—a node at the fixed end and a loop at the lower end. The length of the tube in (a) and the length of the spring in (b) are each a quarter of a wave-length.

If the air column is open at both ends, there must be a loop at each end. Fig. 207 shows one possible mode of vibration. A node is shown at the middle part. Both above and below the node the air moves toward the node and then away from it. This causes relatively large changes in the density of the air at the node. At the loop, where the air has its greatest motion, there is practically no change in the density. This mode of longitudinal vibration of the air is similar to the longitudinal vibration of a rod or a stiff spiral spring clamped at the middle point (sect. 334).

The frequency of vibration of an air column can be computed from equation (6). In a tube closed at one end the length of the tube, L, is one fourth of a wave-length, or

$$\lambda = 4L.$$

Hence, for a closed tube vibrating as in Fig. 206, equation (6) becomes

$$n = \frac{V}{4L}.$$

In a tube open at both ends and vibrating as in Fig. 207,

$$\lambda = 2L$$

and

$$n = \frac{V}{2L}.$$

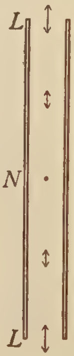


FIG. 207

The cases shown in Figs. 206 and 207 are not the only modes of vibration of air columns. It is not difficult to find others, for it is only to be remembered that there must be (1) a node at a closed end and (2) a loop at an open end. In Fig. 208 are shown a few of the different modes of vibration of both open and closed tubes. The relationships between the length of the tube and the wave-length are given, as are also the relative frequencies of the sounds emitted by a tube vibrating in these modes. The student should verify these relations. To do this it is necessary to remember that the distance between any node and the adjacent loop is one fourth of a wave-length.

In a tube such as is used in a pipe organ or in some other wind instrument, a large number of the possible modes of vibration can be produced simultaneously. In the case of a tube closed at one end, if n_0 is the frequency of the fundamental, the frequencies of the different modes of vibration are n_0, $3n_0$, $5n_0$, $7n_0$, $9n_0$, etc.; hence *the overtones of a tube closed at one end include only the odd harmonics.* In a tube open at both ends the frequencies are n_0, $2n_0$, $3n_0$, $4n_0$, $5n_0$, $6n_0$, $7n_0$, etc.; hence *the overtones of a tube open at both ends include all the harmonics.* Since the quality of a sound depends on the overtones, it follows that the qualities given by the two types of tubes are different.

The material of which tubes are made affects the relative intensity of different overtones and hence affects the tone quality.

In some wind instruments notes of different pitch are obtained by changes in the effective length of the air column. But in a bugle, which has a column of fixed length, only the notes which belong to a harmonic series can be produced, the different modes of vibration being obtained by different ways of blowing.

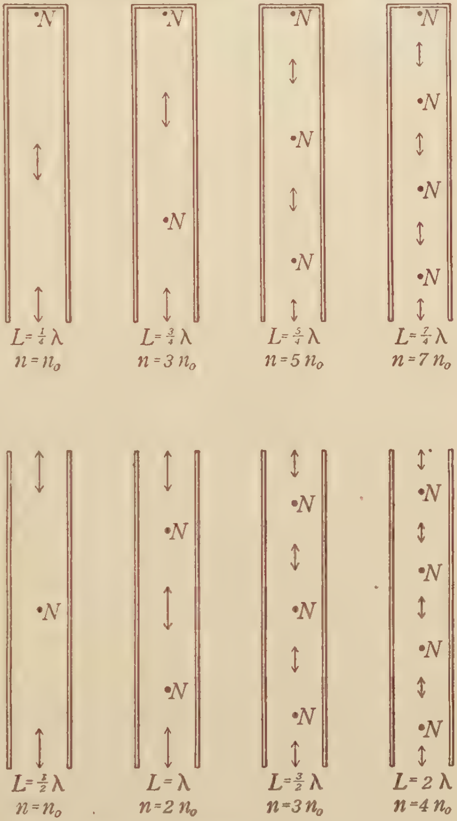


FIG. 208. Modes of vibration of air columns

337. Musical intervals. A musician learns to measure *musical intervals* by his ear, but in many cases a quantitative method of measuring them is necessary. *A musical interval is measured by the ratio of the frequencies of the two notes.* It is usually expressed as the ratio of the larger frequency to the smaller; thus, the interval between two notes of frequencies 500 and 400 is 5/4. The interval is independent of the magnitudes of the frequencies, depending only on their ratio. For example, the interval between notes of frequencies 600 and 400 is the same as that between 1500 and 1000, being 3/2 in both cases.

338. Law of harmonious combinations. *The most harmonious combination of notes is that in which the intervals can be expressed by small whole numbers.*

The interval 2/1 is expressed by the smallest numbers and is the most harmonious combination. It is called the *octave.* The interval 3/2 is next in simplicity. It is called the *major fifth.* The *major fourth* is the interval 4/3. The *major third* is the interval 5/4, and the *minor third* is the interval 6/5.

As the numbers get larger the harmony diminishes. For example, two notes having the interval 9/8 are discordant. This is the interval between *do* and *re* in the major scale.

339. The major scale. The intervals used in the natural major scale are as follows:

do *re* *mi* *fa* *sol* *la* *ti* *do.*
$\frac{9}{8}$ $\frac{10}{9}$ $\frac{16}{15}$ $\frac{9}{8}$ $\frac{10}{9}$ $\frac{9}{8}$ $\frac{16}{15}$

If we arbitrarily take 24 as the frequency of the first *do*, the frequencies of the other notes in the major scale are given in the following table:

do	*re*	*mi*	*fa*	*sol*	*la*	*ti*	do_2
24	27	30	32	36	40	45	48
1							2
2				3			
4		5		6			
	3			4		5	
			2				3
			4		5		6

The ratio numbers in the third and following lines show the harmonious combinations. Thus *do* and *sol* are harmonious because their frequencies are proportional to 2 and 3. Apparently in the formation of this scale the pitch of each note was so chosen that it would harmonize with the others. For example, if we start with *do*, the octave do_2 would be chosen because of the harmony. The next would be *mi* and *sol*, for they are related to *do* by simple ratios. Then *re* and *ti* would be chosen because they harmonize with *sol*, and so on.

PROBLEMS

1. What is the speed of a deep-water wave the wave-length of which is 30 ft.?

2. What is the speed of a water-wave of long wave-length if the water is only 1 ft. deep?

3. If the velocity of a sound-wave is 330 m./sec., and the wave-length is 66 cm., what is the frequency?

4. (*a*) Compute the wave-length of a sound-wave produced by a high soprano singing a note of pitch 1300. (*b*) Compute the wave-length of the low bass pitch 60. (The temperature is 20° C.)

5. Thunder was heard 5 sec. after a flash of lightning was seen. How many feet away was the lightning? (Assume the temperature of the air to be 20° C. Do not take into account the speed of light.)

6. A gun is fired, and in 3.2 sec. the echo is heard. How far away is the cliff which reflects the sound? (The temperature is 0° C.)

7. A whistle has a frequency of 300 vibrations per second. What will be the apparent frequency if it is moving away from the observer with a speed of 88 ft./sec. (60 mi./hr.)? (The temperature is 20° C.)

8. Compute the musical interval between the two tones heard, one just before and the other just after a whistling engine has passed at 60 mi./hr. (Assume the temperature of the air to be 0° C.)

9. An open organ-pipe has a length of 40 cm. What will be the frequency of the fundamental at 20° C.? of the first overtone?

10. A tube, closed at one end, is 50 cm. long. What will be the frequency of vibration, at 20° C., of its fundamental mode? of its first overtone? of its second overtone?

11. Compute the frequency of the fundamental note of a child's tin whistle, consisting of a tube 3 cm. long, closed at one end, when blown with air which has a temperature of 30° C.

12. A tube 42 cm. long, closed at one end, is in resonance with a tuning-fork. What is the number of vibrations per second of the fork? (The temperature is 20° C.)

13. In an experiment with the dust figures of a Kundt's tube, at 20° C., it was found that a rod 60 cm. long, clamped at its center, produced air-waves which were 8.2 cm. long. (*a*) What was the frequency of vibration? (*b*) What is the velocity of a compressional wave in the rod?

14. At international pitch "violin *A*" has a frequency of 435 vibrations per second. If *C* is taken as the *do* of a major scale, compute the frequency of *C*, *E*, and *G* of the octave containing this *A*.

PART IV. MAGNETISM AND ELECTRICITY

CHAPTER XXV

MAGNETISM

Magnets, 340. Magnetic forces act through most substances, 341. Poles, 342. Poles always exist in pairs, 343. Attractive forces have a limited range, 344. The magnetic compass, 345. Two kinds of poles, 346. The magnetic field, 347. Magnetic lines, 348. The earth a great magnet, 349. Declination, 350. Inclination, or dip, 351. Variations in the earth's magnetic field, 352. Importance of a knowledge of the earth's magnetism, 353. Magnetic induction, 354. The tendency of soft iron to move into stronger fields, 355. Retention of induced magnetization, 356. An explanation of the behavior of iron filings in a magnetic field, 357. Other magnetic substances, 358. Effect of changes in temperature, 359. Magnetism and the molecular theory, 360. The common electrical units are based upon magnetic measurements, 361. The quantitative law of attraction between two poles, 362. Strength, or intensity, of a magnetic field, 363. Strength of a magnetic field a vector quantity, 364. Use of magnetic lines to indicate strength of a field, 365. Magnetic moment, 366. Energy in a magnetic field, 367. Summary, 368.

340. Magnets. That magnets in the form of bars of hardened steel exert attractive forces on pieces of iron and steel is well known. Pieces of iron ore that are already magnetized are sometimes found. These are called *natural magnets* or *lodestones.* Magnetic attraction has been a familiar phenomenon a very long time,—how long we do not know. Certainly Thales (about 600 B.C.), one of the "seven wise men" of early Greece, knew something about it. Some very old tales refer to magnetic action. For example, in one of the Arabian Nights tales it is stated that there was a mountain magnetized so strongly that ships sailing near it were wrecked because the nails were pulled out of the timbers.

Only a few substances are noticeably attracted by magnets, the most important being iron, steel, nickel, and cobalt.

341. Magnetic forces act through most substances. It is very easily shown that small pieces of iron—tacks, iron filings, etc.—are attracted by a magnet, even though a sheet of glass or a board is placed between the magnet and the pieces of iron. Trial with other substances shows that the magnetic attraction takes place through all substances, with no appreciable change except through those that are attracted by the magnet. A sheet of iron acts as a partial screen; a thin sheet produces a small effect, but thick pieces, especially when they completely inclose a region, greatly diminish the magnetic attraction.

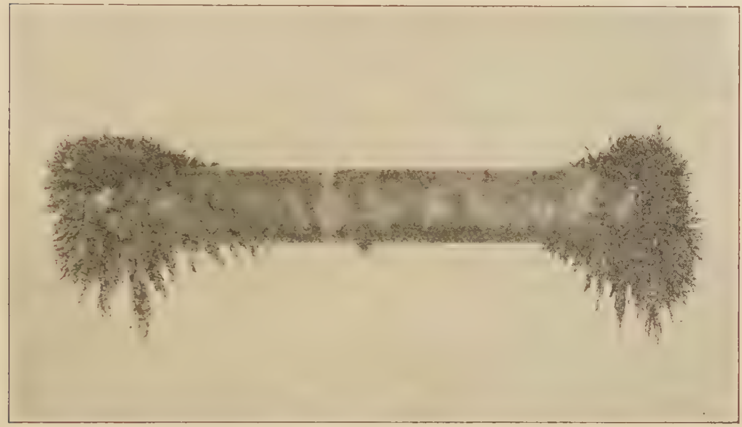


FIG. 209. Iron filings clinging to a magnet

342. Poles. A magnet exerts the greatest attractive force at points near the ends. The part halfway between the ends shows very little attractive force, often none. This is very clearly shown if a magnet is thrust into iron filings and then lifted out. The filings will be found sticking to the magnet in large clusters around the ends (Fig. 209). On this account it was once thought that the "seat" of attraction was in or near the ends. These apparent seats of attraction were called *poles*. Every bar magnet has at least two poles, usually one near each end.

343. Poles always exist in pairs. It has not been found possible to make a magnet which has only one pole. If one takes a piece of tool steel (such as a knitting-needle or a piece of hack-saw blade) that has been hardened until it is brittle, and magnetizes it, he will find on breaking it up into pieces that each piece will be a complete magnet with two poles.

344. Attractive forces have a limited range. The attractive or repulsive force exerted by one magnet on another or by a magnet on iron is never large unless the two are close together.* Even the powerful lifting-magnets used in handling large masses of iron exert little effect until the iron to be lifted is very close to the poles.

345. The magnetic compass. If a bar magnet is suspended so that it is free to rotate about a vertical axis, it will finally come to rest, pointing in a definite direction. If there are no other magnets or masses of iron near, it will point approximately north and south. The magnetic *compass* usually consists of a light magnetized steel needle mounted on a sharp point, about which it can turn with very little friction. No one knows who discovered and first utilized a magnetic compass. The Chinese knew about it as long ago as A.D. 121, but how this knowledge was brought to Europe we do not know. Sometime during the tenth or eleventh century the tendency of a magnet to point north and south became known, and in the twelfth century, or perhaps earlier, it was used in navigation.

The pole that points north is called the *north-seeking* pole, and that pointing south the *south-seeking* pole.

346. Two kinds of poles. It is always the same pole of a magnet that tends to point northward. This shows that the two poles of a magnet are not alike. Further confirmation of this difference can be obtained by the following simple experiment: A bar magnet in a stirrup is suspended by a thread from some suitable support, or a moderately long compass needle may be used. When a second magnet is brought near the suspended one, the forces of attraction and repulsion are easily observed.

---

* The force between two magnets at distances much larger than the lengths of the magnets varies inversely as the fourth power of the distance,—a very rapid decrease in force as the distance increases.

1. One of the poles of the second magnet will repel one of the poles of the suspended magnet, but will attract the other. This shows that the two poles of the suspended magnet are not alike.

2. If the other pole of the second magnet is now used, it also will attract one pole of the suspended magnet and repel the other. But the pole now attracted is the one that was repelled in the first case, and the pole now repelled is the one that was attracted in the first case. This shows that the two poles of the second magnet (the one held in the hand) are not alike.

If the poles of other magnets are brought near the suspended magnet, it will be found that one of the poles of each magnet attracts the north-seeking pole and repels the south-seeking pole of the suspended magnet, while the other pole attracts the south-seeking pole and repels the north-seeking pole of the suspended magnet. In this way it can be shown that there are two kinds of poles and that every magnet always has one of each kind.

The following conclusion can be drawn from experiments similar to those just described: *Like poles repel and unlike attract.*

347. The magnetic field. The region around a magnet where a force would act on a small piece of iron or on another magnet brought into that region is usually called the *field* of that magnet. If a short compass is brought into this field, it will at each place point in a definite direction. This direction can be approximately predicted by the following method: In Fig. 210 the vector OA is drawn* to represent the force produced by the north-seeking pole of the magnet on the north-seeking pole of the compass. The vector OB represents the attractive force the south-seeking pole of the magnet has for the north-seeking pole of the compass. The direction of the resultant force is given by OR. The force acting on the south-seeking pole of the compass will be just opposite to OR. The direction of the line OR will be the direction that the compass needle points.

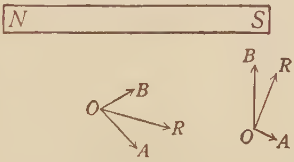


FIG. 210

---

* Diagrams are drawn at different places to show that the result is different in different parts of the field.

The direction in which the north-seeking pole of the compass needle points is called the direction of the magnetic field at that place.

It is a very instructive experiment to map out, with the aid of a small compass, the direction of the magnetic field at a large number of points near a magnet.

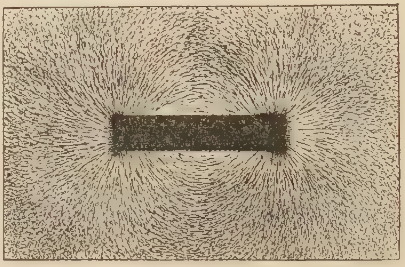


FIG. 211

348. Magnetic lines. The following experiment should be seen by each student. A sheet of glass or smooth cardboard is laid on a bar magnet, and iron filings are sprinkled evenly over the glass. If the glass is tapped or slightly jarred, the filings will arrange themselves in lines which run from one pole around to the other. This is shown in the photographic reproduction given in Fig. 211.

It is an interesting fact that these lines always take at each point the direction of the magnetic field. Hence a magnetic field can be mapped out either by observing at a large number of places the direction a compass points or by using the lines formed by iron filings.

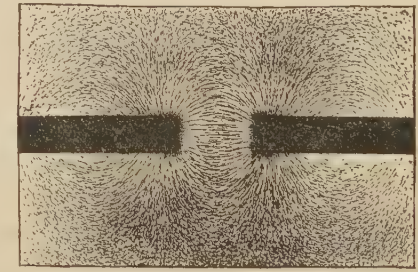


FIG. 212

Fig. 212 shows the lines traced out by filings in a field between two unlike poles; Fig. 213 is for the case of two like poles.

If a series of lines are drawn through the field in such a way that at each point of the field the direction of the line through that point is in the direction a small compass points, they are called *magnetic lines*.

It is customary to say that magnetic lines always run from a north-seeking pole to a south-seeking one, in the direction a compass points. Hence *magnetic lines give at each point the* *direction of the force that would act on a north-seeking pole if such a pole were placed at that point.*

The attraction or repulsion between magnetic poles can be explained if we imagine the magnetic lines to have the following properties: (1) they are under tension and therefore tend to contract; (2) they repel each other sidewise.

If the lines are under tension and trying to shorten, the attraction between the unlike poles in Fig. 212 would be explained. As shown in the figures, the lines often take an indirect path from one pole to the other. This bulging of the lines is imagined to be due to side repulsion of the lines for each other. Any line, because of its tension, tries to move in between the poles and take the shortest path. But the repulsion of the lines inside holds it out, in the position shown. The distortion of the lines in Fig. 213 is also due to this repulsion. This gives a sort of visual explanation of how the two poles of Fig. 213 repel each other.

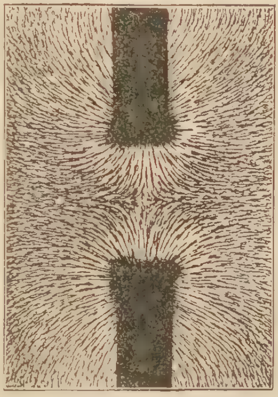


FIG. 213

It should be understood that these magnetic lines are purely imaginary. But they are so very useful for purposes of study that many electricians are in the habit of referring to them as if they actually existed in the space around every magnet.

349. The earth a great magnet. It seems that an Englishman, William Gilbert, in the year 1600, was the first to point out that the earth is a great magnet, and that this is the reason why a magnet always tends to point in an approximately north-and-south direction. The magnetic poles of the earth do not coincide with the geographical poles. If we call the line joining the two poles of a magnet the *magnetic axis* of the magnet, the magnetic axis of the earth does not coincide with its axis of rotation. If the magnetic axis is prolonged until it cuts through the surface, it emerges in the Northern Hemisphere about one thousand miles from the north geographical pole, just inside the arctic circle and about the longitude of Omaha and Winnipeg. The other end of this line, extended out through the surface, would emerge south of Australia.

We have called the pole of the compass which points toward the north the *north-seeking pole.* Since this is attracted by the magnetic pole of the earth which lies to the northward, the arctic pole must be a *south-seeking* one, and the antarctic magnetic pole a *north-seeking* one. The student can see at once the confusion which would follow if we called the poles of a magnet the *north* and *south* poles. It is to avoid such confusion that many prefer to call the *north-seeking* pole the *positive* pole, and the *south-seeking* the *negative* pole. In these terms the arctic magnetic pole is a negative and the antarctic a positive pole.

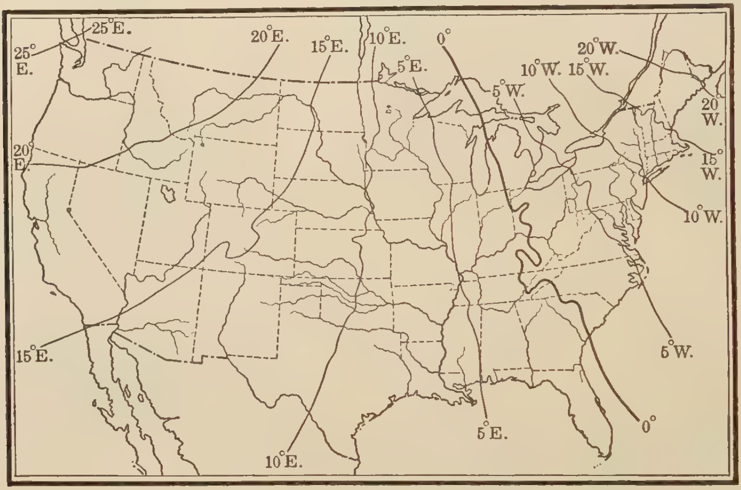


Fig. 214. Lines of equal magnetic declination

350. Declination. In only very limited regions of the earth does the compass point exactly north and south. The angular distance the compass points away from true north is called the *declination.* Columbus used a compass on his voyage to America. Indeed, it was on this voyage that the sailors made what was to them a very startling discovery: that the declination was not the same everywhere but changed as they went westward. Fig. 214 shows the declination throughout the United States. The irregular lines running across this map run through points having the same declination. Such lines are called *isogonic* lines. The direction of a line at any locality does *not* give the direction the compass needle points. This direction is marked near the end of each line. At points on a line running through Michigan, Ohio, eastern Kentucky, and Tennessee, and then through South Carolina, the compass points directly north. At all localities east of this line the compass points west of north; at places west of this line it points east of north. The extreme variations in the United States are found in Maine and in Washington. In northeastern Maine the compass points 20 degrees west of north; in the Puget Sound region it points nearly 25 degrees east of north.

351. Inclination, or dip. If a magnetic needle is balanced accurately before being magnetized and is then mounted so that it is free to turn in any direction, it will not come to rest in a horizontal position. In the Northern Hemisphere the north-seeking pole dips downward; in regions on a line running through Maine, Ontario, central Michigan, and Minneapolis, and then to the northwest, it dips about 75 degrees; along a line running near Washington, D.C., and St. Louis, and then northwest to a point near Seattle, it dips 70 degrees; along a line running near Charleston, South Carolina, and Little Rock, Arkansas, and thence northwest through Nevada and northern California, it dips 65 degrees. In general, as one goes northward the dip increases until it becomes 90 degrees over the arctic magnetic pole; as one goes southward the dip gets less, gradually becoming zero; still farther south the other pole dips. This angular departure from the horizontal is called the *inclination* or, more simply, the *dip*.

The earth's magnetic lines in the United States must be thought of as running nearly northward and obliquely downward.

352. Variations in the earth's magnetic field. The magnetic condition of the earth is slowly changing. At present both the declination and the inclination are increasing in the greater part of the United States. This change is exceedingly slow, being at no place greater than about six minutes per year, and through most of the country less than that. These slow changes are called *secular* changes.

The *diurnal* variation is the one that takes place daily. Both the inclination and the declination change slightly with the time of day. The greatest daily variation in the United States is not greater than an angle of fifteen minutes, and in some localities it is as low as four minutes.

In addition to these changes there are *irregular* disturbances, which sometimes produce rather large changes in the magnetism of the earth. These are often called *magnetic storms*. In some of these storms a sensitive compass needle will swing to and fro, departing two or three degrees from its normal position. The cause of these storms is not understood, but they are often accompanied by sun spots and auroral displays. During a magnetic storm currents of electricity in the earth are produced which may put long-distance telephone and telegraph lines out of use for several hours.

353. Importance of a knowledge of the earth's magnetism. In the great majority of cases ships are navigated by the aid of the compass. Even along our coasts and on the Great Lakes the compass is indispensable. Therefore it is highly important that the value of the magnetic declination be known and charted for all the navigable parts of the waters of the earth. The Carnegie Institution has been notable in this work. In 1905 the wooden brigantine *Galilee* was fitted up for explorations of the Pacific. After three years of work the Pacific Ocean routes were well charted. Although the *Galilee* was built of wood and was a sailing-vessel, its "magnetic constants" were large enough to necessitate many corrections of the observations. As a result, in 1909 the nonmagnetic yacht *Carnegie* was built. In the construction of this craft practically no iron was used: bolts of copper and bronze replaced those of iron. Even the gas engine was built almost entirely of bronze. Since that time this ship has been used a great deal, and by its aid valuable observations have been made over large portions of the international highways.

354. Magnetic induction. If a piece of soft iron, not magnetized, is brought near the pole of a bar magnet, the soft iron becomes magnetized. This can easily be shown by the tendency of tacks or iron filings to cling to the soft iron (Fig. 215). On removing the bar magnet the iron at once loses its magnetism. This method of magnetizing a piece of iron or steel is called magnetization *by induction.*

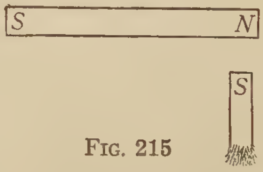


FIG. 215

Notice in Fig. 215 the kinds of poles induced in the soft iron. An unlike pole is produced near the north-seeking pole of the bar magnet, while the like pole is produced at the farther end of the soft iron.

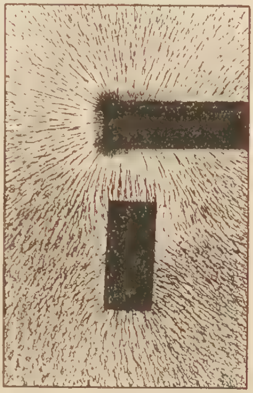


FIG. 216

Fig. 216 is reproduced from an actual experiment using iron filings. The smaller outline was a piece of soft iron; the larger, a bar magnet. Apparently there is a tendency for magnetic lines to crowd into the soft iron rather than go through the air. This same thing is shown when an iron washer is placed in a magnetic field (Fig. 217). Note the almost entire absence of lines inside the ring. The magnetic lines, preferring the iron path, leave only a very weak field in the center. The space in a cavity in a large mass of iron is almost entirely screened off from any magnetic influence outside the iron. A compass placed in such a cavity would show very little tendency, if any, to point northward.

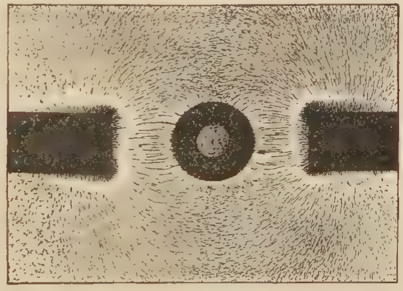


FIG. 217

An interesting experiment showing magnetic induction is the following: A piece of iron pipe three or four feet long is held in the direction of the earth's magnetic field; that is, in a north-and-south direction, but dipping down about 70 degrees. After the pipe is jarred by being hit with some hard object, it will be found, on testing it with a compass needle, to be magnetized, having at the lower end a north-seeking pole (one unlike the arctic magnetic pole toward which it is pointing) and at the upper end a south-seeking pole. If the pipe is inverted and jarred, the magnetism in it will be reversed; the lower end always becomes a north-seeking pole.

Nearly all pieces of iron that stand continually in some definite position become magnetized by induction from the earth's field.

Iron hitching-posts and fence posts show by their action on a compass needle that they are feebly magnetized. In a house the water, gas, and steam pipes, the stoves, and the stovepipes are usually magnetized.

The student will learn that all methods of magnetizing iron or steel are cases of magnetization by induction.

355. The tendency of soft iron to move into stronger fields. If the pole of a magnet is brought near a small piece of iron, the piece of iron may jump up and cling to the pole. Why? It is not sufficient to say that the piece of iron is attracted. The piece of iron becomes magnetized by induction and acquires two equal poles. One of these poles is attracted, the other is repelled. *It is only when the attractive force is greater than the force of repulsion* that the resultant force acting on the piece of iron is not zero.

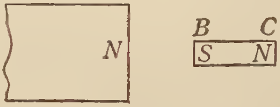


FIG. 218

Whenever a piece of iron is placed in a uniform field, the forces acting on the two induced poles are equal and opposite; but when the field is not uniform, the forces may not be equal and opposite. Figs. 218 and 219 give two cases where the forces on the two induced poles are not equal and opposite. Fig. 218 shows a pole of a strong magnet placed near the soft iron BC. Since B is nearer the magnet than C, the force of attraction on B is greater than the repulsion of C. In Fig. 219, A and D are two strong poles. BC is soft iron. There are really two forces acting on each of the induced poles of BC. The arrows show only the stronger of each force. It is obvious that the resultant force on BC is not zero. The soft iron is pulled upward into the strong field which lies between the two poles A and D. The student can find many other cases. The magnetic field of the earth gives at any one locality a very good example of a uniform field. It produces opposite and practically equal forces on the two poles of a compass needle or other magnet, with the result that there is practically no force of translation due to the earth's field.

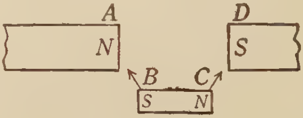


FIG. 219

356. Retention of induced magnetization. Soft iron or steel does not retain magnetism very well. Short pieces lose practically all of it when they are removed from the magnetizing field. Large masses of iron or long pieces of iron or steel retain magnetism better. Hardening the steel increases its ability to retain magnetism.

To make a permanent magnet, "tool" steel should be used, and it should be made "glass" hard—so hard that it becomes brittle. This can be done by heating it to a cherry red and then quickly plunging it into water. The steel bar is then placed in a strong magnetic field, and magnetized by induction.

357. An explanation of the behavior of iron filings in a magnetic field. When iron filings are sprinkled on a plate in a magnetic field, each small particle becomes magnetized by induction. Each particle has then two poles—one north-seeking and the other south-seeking. These little magnets not only attract and repel each other but act like little compass needles. When the plate is jarred so that they can move, they tend to turn and set themselves parallel to the field. The student should be able with these suggestions to explain why the filings tend to group themselves into strings.

358. Other magnetic substances. A piece of nickel or cobalt will cling to the pole of a magnet. In other words, nickel and cobalt become magnetized by induction. Of all metals, only these two are comparable to iron and steel in their magnetic properties, and even they are far below iron in this respect. Certain alloys, such as an alloy of manganese, aluminum, and copper, or of manganese and bismuth, are magnetic, even though the constituents taken individually are not.

A number of substances are feebly attracted by very powerful magnets: certain of the rare earths, paper, porcelain, tungsten, platinum, and aluminum. Copper, bismuth, antimony, and many other substances are feebly *repelled* from the poles of a very strong electromagnet (see section 501).

359. Effect of changes in temperature. When iron or steel is gradually heated, it suddenly loses almost all its magnetic properties at a dull-red heat (a temperature between 700° and 800° C.). A steel magnet raised above this temperature will, when cooled, be found to have lost all its magnetism. It can be remagnetized, however. Warming a permanent magnet causes a temporary decrease in its magnetism, while cooling it produces a temporary increase.

360. Magnetism and the molecular theory. A simple explanation of magnetic action is based on the supposition that iron is made up of small parts, each of which is permanently a complete magnet.

As usually stated, it is the molecule that is this complete tiny magnet. (There are good reasons for thinking that it is the atom.)

According to this theory, in an unmagnetized piece of iron the molecules are turned in all sorts of directions. Being magnets they will attract and repel each other, and there will be a tendency for them to turn so as to bring their unlike poles near together. Two possible simple groupings of small bar magnets are shown in Fig. 220. With the molecules of a bar of iron arranged in many such groups with unlike poles together, there will be no appreciable external field.

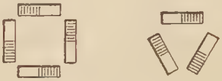


FIG. 220

When the iron is partly magnetized, some of these groups are broken up, and some such arrangement as is shown (on an exaggerated scale) in Fig. 221 might exist. When completely magnetized *all* the tiny magnets will be pointing in the same direction, as shown in Fig. 222.

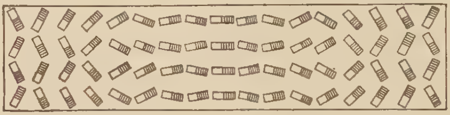


FIG. 221. A partially magnetized iron bar

An important fact about magnets that needs explanation is the apparent concentration of attractive and repulsive forces at the ends, the central parts of a magnet showing very little if any force. This is easily explained by the aid of Fig. 222. Along the central part of the magnet unlike poles are close together, one north-seeking pole completely neutralizing the effect of the adjacent south-seeking pole, as far as distant points are concerned. Hence the central parts would neither attract nor repel a pole placed near the magnet. Near one end there is an excess of north-seeking poles, and at the other end an excess of south-seeking ones. Consequently one end of the bar should attract a south-seeking pole of another magnet and repel a north-seeking one.

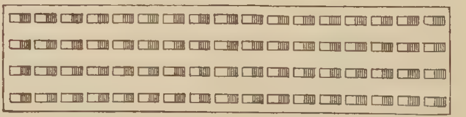


FIG. 222. A completely magnetized iron bar

An explanation of the fact that each part of a broken magnet is a complete magnet is obvious from this theory.

Mention has been made of the fact that jarring will assist in the magnetization of a bar of iron or steel. This is explained by the assumption that jarring assists the turning of the molecules. When the molecules of a magnet are lined up as shown in Fig. 222, they are not in equilibrium as regards their mutual magnetic forces. When the magnetizing agency is removed, the molecules have a tendency to turn around and form small groups. Hardening the steel apparently tends to prevent them from turning, but jarring will assist them and thus tend to weaken their magnetic action. Hence magnets must be handled carefully.

361. The common electrical units are based upon magnetic measurements. As soon as commercial applications of electricity began, there arose a demand for accurate methods of measuring all electrical quantities. Fortunately the fundamental laws of magnetism and electricity are in general not rough approximations, but are correct to a very high degree of accuracy. The development of ingenious instruments has brought electrical measurements to a degree of accuracy that is surpassed in only a few cases in all physical science.

The common commercial units in electricity are based either directly or indirectly on magnetic quantities. In order to understand the commonly used electrical methods, certain rather simple magnetic units and laws must be understood. These will be given in the following sections.

362. The quantitative law of attraction between two poles. By using two long thin magnets with two poles near each other, and the other poles so far away that they exerted practically no effect, Coulomb (1736–1806) was able to determine experimentally the law of force between two poles. He found that the attraction or repulsion between two poles varies inversely as the square of the distance between them* and that the attraction or repulsion between two poles is proportional to the product of the strength of the poles.

These facts are best expressed in the following form:

$$\text{Force of attraction or repulsion} = k\frac{m_1 m_2}{d^2}, \tag{1}$$

---

*It should be noted that this statement applies to the attraction between two *poles*. The footnote on page 362 refers to the attraction between two *magnets*.

where m_1 and m_2 are the strengths of the two poles, and d the distance between them. The value of the constant k, as explained in section 65, depends on the units used in measuring force and the quantities m and d. In the system known as the *C.G.S. electromagnetic system* the value of the unit for measuring the pole strength is so chosen that k is equal to unity when the space between the poles contains air or some other nonmagnetic medium.

Hence, in C.G.S. units, $$F = \frac{m_1 m_2}{d^2}, \qquad (2)$$

where F is the force in dynes and d the distance in *centimeters.*

From this equation it appears that if m_1 and m_2 are each equal to unity, and if they are placed so that d is unity, then the force F is unity. Hence we define our *unit* pole as follows: *A unit magnetic pole when placed 1 centimeter from another unit like pole in air will repel it with a force of 1 dyne.*

Careful experiments show that the two poles of a magnet always have equal strengths. If the north-seeking end has a strength of 40, the other pole will also have a strength of 40. Often these strengths are indicated by $+40$ and -40. If these signs are applied in equation (2), a positive sign of the product on the right-hand side means a repulsion, and a negative sign an attraction.

363. Strength, or intensity, of a magnetic field. In many different cases in electrical work it is necessary to know something about the strength of magnetic fields. By general agreement it has been decided that *the strength, or intensity, of a magnetic field at any point shall be measured by the force the field exerts on a unit north-seeking pole placed at that point.*

This is the *quantitative* definition of strength, or intensity, of a magnetic field.

If H represents the strength of a certain field, and m the strength of a pole which has been brought into this field, the force, F, acting on this pole is given by the equation

$$F = Hm. \qquad (3)$$

This follows from the definitions; for H is the force on a unit pole, and m the number of units in the pole.

364. Strength of a magnetic field a vector quantity. Since the strength of a magnetic field is equal to the force on a unit north-seeking pole, it has *direction* as well as *magnitude*; hence it is a *vector* quantity. In section 347 it was shown that vector methods could be used to get the direction a compass needle will point when brought near a magnet. The same method will give not only the direction of the strength of the field but also its magnitude if the vectors *OA* and *OB* in Fig. 210 represent the forces acting on a *unit* pole placed at *O*.

The strength of the earth's magnetic field varies from place to place on the earth's surface. In the United States it has a value of about 0.6 C.G.S. unit, and its direction makes an angle of about 70 degrees with the horizontal (the student should understand that when one states that the strength is 0.6, he means that the force the earth's field exerts on a unit pole is 0.6 dyne). In Fig. 223, *OB* represents the strength of the earth's field, *OA* the *horizontal* component, and *OC* the *vertical* component. Since compass needles, and the magnets in many instruments, rotate about a vertical axis, it is only the horizontal component of the earth's field that is effective in turning these magnets around. The horizontal component varies from 0.15 C.G.S. unit, along a line running nearly through Duluth, Ottawa, and Montreal, to 0.27 C.G.S. unit in the extreme southern part of the United States. Why should it be greater in the southern part of the United States?

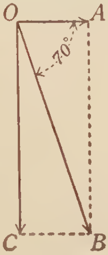


FIG. 223

365. Use of magnetic lines to indicate strength of a field. The strength, or intensity, of a magnetic field is often represented by the way the *magnetic lines* are drawn. It has already been pointed out that the direction of these lines is identical with the direction of the field. It is customary to make the density of the lines in air (that is, the number passing through an area of 1 square centimeter taken perpendicularly to the lines) numerically equal to the strength of the field. So common is this practice that when the strength of a field is to be stated, many say not that the strength of the field is, for example, 50 C.G.S. units, but rather that the number of magnetic lines per square centimeter is 50.

366. Magnetic moment. The magnetic moment of a magnet is defined as *the product of the strength of one of the poles times the distance between the two poles of the magnet.* Or

$$M = ml, \tag{4}$$

where M is the magnetic moment, m the strength of one of the poles, and l the distance between the poles.

Imagine a bar magnet placed at right angles to a field the intensity of which is H (Fig. 224). The force on each pole is mH, according to equation (3). Since these forces are equal but in opposite directions, they form a "couple" (sect. 14) and tend to rotate the magnet. The torque is equal to the product of one of the forces times l, the distance between the poles. Or

$$\begin{aligned}\text{Torque} &= mHl \\ &= MH, \qquad (5)\end{aligned}$$

since M, the magnetic moment, is equal to ml. It is not very difficult to measure the torque required to hold a magnet perpendicular to a field. If the intensity of the field, H, is known, M can be calculated; or, if M is known, H can be determined.

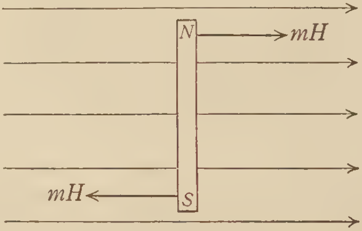


FIG. 224

In practically all cases of forces between magnets (complete magnets, not poles) the *magnetic moment* appears in the algebraic expression for the force or for the torque. It is also true that in nearly all experimental methods of measuring the strength of magnets it is the magnetic moment that is determined. Hence in magnetic measurements magnetic moment is an important quantity.

367. Energy in a magnetic field. It is always true that work must be done to create a magnetic field. There is good evidence for believing that a magnetic field has energy. It can be shown (though the details are beyond the scope of this book) that the energy per cubic centimeter of a magnetic field in air is

$$\frac{H^2}{8\pi}, \tag{6}$$

where H is the strength of the field.

368. Summary. Magnets will attract pieces of iron, nickel, and cobalt, and, very feebly, some other substances. The ends where the attractive force is the greatest are called *poles*. There are two kinds of poles, *north-seeking* and *south-seeking*. Each magnet has at least one of each kind. Like poles repel and unlike attract.

Magnetic lines are imaginary lines the *direction* of which gives at any one point the direction of the force in air on a north-seeking pole placed at that point. A short compass needle will always set itself tangentially to a magnetic line.

The *declination* of the earth's magnetic field is measured by the angular deviation of the compass from the north. It is different in different parts of the United States.

The *inclination*, or *dip*, is the angle the magnetic lines of the earth make with the horizontal. It varies from 65 to 75 degrees in different parts of the United States.

The earth's magnetic field is not constant but is gradually changing. In addition to the secular changes, there are daily changes.

When a piece of soft iron is brought into a magnetic field, it is magnetized by *induction*. The magnetic lines of the field tend to crowd into the iron rather than to go through the air. The attraction of pieces of iron by magnets can be explained by the fact that the pieces of iron become magnetized. Hardened steel does not readily become magnetized, but it retains part of the magnetism when carried out of the field. It is because each particle of iron becomes a complete magnet that iron filings will arrange themselves in lines which in general are in the same direction as the magnetic lines of the field.

Iron loses its magnetic properties at a temperature between 700° C. and 800° C.—a dull-red heat.

All the simple facts of magnetism can be explained by the molecular theory, which assumes that each molecule is a complete magnet.

Coulomb's quantitative law of the force acting on each one of two poles in air is, by equation (1),

$$F = k\frac{m_1 m_2}{d^2}.$$

In the C. G. S. electromagnetic system of units the unit for measuring m, the pole strength, is so chosen that k is equal to unity when the poles are separated by air.

A C. G. S. unit north-seeking pole, when placed 1 centimeter away from an equal north-seeking pole in air, will repel it with a force of 1 dyne.

The force acting on a pole in a field the strength of which is H is, by equation (3),

$$F = Hm.$$

The *strength* of a magnetic field is equal in both direction and magnitude to the force exerted on a unit north-seeking pole placed in that field.

It is customary to make the density, or the number per square centimeter, of magnetic lines in air numerically equal to the strength of the field. This is entirely conventional.

The magnetic moment of a bar magnet is the product of the strength of one pole by the distance between the poles, or it is equal to the torque necessary to hold the magnet at right angles to a magnetic field of unit strength. The poles of any magnet always have equal pole strengths.

PROBLEMS

1. Find the force of attraction between two poles $+30$ and -20 when placed 10 cm. apart. Find the force when they are 20 cm. apart.

2. Two equal poles attract each other with a force of 64 dynes when placed 5 cm. apart. Find the strength of one of the poles.

3. (*a*) Find the strength of a magnetic field at a point halfway between two poles, $+30$ and -20, when the poles are 20 cm. apart. (*b*) Find the strength of the magnetic field when the second pole is replaced by one of strength $+20$.

4. The poles of a bar magnet are 10 cm. apart. The strength of each pole is 30 C.G.S. units. Compute the strength of the magnetic field at a point 10 cm. from one pole and in the direction of the prolongation of a line joining the two poles.

5. A magnet has a magnetic moment of 600 C.G.S. units. What are the strengths of the poles if they are 20 cm. apart?

6. A magnet has poles of strength 200 C.G.S. units which are 15 cm. apart. Find the strength of the field at a point 15 cm. from each pole.

7. A pole of strength 20 C.G.S. units is placed in a magnetic field due to another magnet. The force acting on the pole is 70 dynes. Find the strength of the field which is produced at this place by the other magnet.

8. A magnet, magnetic moment 1200 C.G.S. units and length 20 cm., is placed at right angles to the earth's magnetic field. It is found that the force on one of the poles is 36 dynes. What is the strength of the earth's magnetic field?

9. A magnet 20 cm. long, magnetic moment 800 C.G.S. units, is held in a uniform magnetic field at right angles to the direction of the field. The strength of the field is 6 C.G.S. units. Find (*a*) the force on each pole and (*b*) the torque acting on the magnet.

CHAPTER XXVI

ELECTROSTATICS

Introduction, 369. Electrified bodies, 370. Two states of electrification, 371. Conductors and nonconductors, 372. Charges of electricity, 373. Fluid theories of electricity, 374. The negative-electron and positive-nucleus theory, 375. The leaf electroscope, 376. Both kinds of electricity produced simultaneously, 377. Electrostatic induction, 378. An explanation of the action of the electroscope, 379. Charging an electroscope by induction, 380. The attraction of pith balls and other bodies explained, 381. The return shock, 382. The electrophorus, 383. Electric charges from dry cells, 384. Electric charges from a power circuit, 385. The distribution of charges on conductors, 386. The action of points, 387. The lightning rod, 388. The law of force between two small electrified spheres, 389. Unit charge, 390. A natural unit of electric charge, 391. The intensity, or strength, of an electric field, 392. Electrical lines: the strain theory, 393. Faraday's ice-pail experiment, 394. The Leyden jar, 395. A brief history of early electrical discoveries, 396. Summary, 397.

369. Introduction. It is convenient to discuss the properties of electricity under two heads: electricity at rest (*electrostatics*) and electricity in motion (*electric currents*). When electricity is at rest, it has certain properties; when it is in motion, it has certain *additional* ones. It happens that the most convenient sources of electricity for showing the electrostatic properties are different from those used in getting electric currents. This should not lead the student to believe that he is dealing with different things.

370. Electrified bodies. The fact that certain bodies—amber, gums, etc.—when rubbed with a cloth acquire the property of attracting light objects has been known a long time, how long we do not know. In the dry weather of deep winter the electrification of fur, wool, silk, and paper is commonly observed. Children learn that sometimes in cold weather they become so electrified by shuffling over a rug or carpet that a vigorous spark occurs when they touch another person.

Hard-rubber rods, sealing-wax, celluloid, dry paper, and sulfur, when rubbed with a woolen cloth or on a woolen sleeve, acquire the property of attracting light bodies. Hard-glass tubing when thoroughly dry is also easily electrified. The list is almost endless. Practically every solid, when rubbed with a different substance, acquires this property. Bodies may be electrified by liquids; gasoline, for example, electrifies pipes through which it flows. In dry-cleaning establishments the electrification of woolens and silks when being washed with gasoline is a source of danger, for a small spark may ignite the gasoline. If one is watching for it, one may, in cold, dry weather, find many times a day different objects that have become electrified.

When electrified, a body has the property of attracting not only light bodies, such as small pieces of paper, feathers, and pith, but also large bodies. If the larger bodies are free to move so that a small force will move them, this attraction can be easily shown. For example, a horizontal meter stick suspended in a wire stirrup by a thread can easily be made to move by the attraction of an electrified body. Light bodies are commonly used to show this attraction merely because the forces involved are small.

The forces produced by electrified bodies are not a special type of force. They obey the same laws of mechanics as all other forces; for example, they obey Newton's laws of motion. Not only will an electrified rod hung in a stirrup attract another object, but the other object (for instance, the hand) will attract it. This is an example of action and reaction. If the rod attracts the hand, the hand must attract the rod.

371. Two states of electrification. If one balances a hard-rubber rod, which has been electrified by rubbing it with wool, in a stirrup hung by a thread, so that the rod is free to swing around, and then brings near it another similarly electrified rubber rod, he will find that they *repel* each other. Similarly, a piece of sealing-wax electrified by rubbing it on wool will repel the electrified hard-rubber rod. Further testing will show that while an electrified hard-rubber rod repels another hard-rubber rod, a rod of hard glass when rubbed with silk will *attract* the electrified hard-rubber rod. From this one concludes that there must be more than one state of electrification, and that *like electrifications repel each other and unlike attract*. Further testing shows that this is a general rule.

Positive is the name given to the kind of electricity found on a hard-glass rod which has been rubbed with silk, and *negative* to the kind found on hard rubber when it has been rubbed with wool.

It is rather difficult to get good results with a glass rod, except in very dry weather, unless it is kept hot. On the other hand, hard rubber is easily electrified under nearly all atmospheric conditions. The imitations of hard rubber which are on the market do not all behave like hard rubber; some of them will give positive electrification when rubbed with wool.

372. Conductors and nonconductors. There are *no conductors of magnetism*. Touching a magnet with the hand does not take away the magnetism, nor can magnetism be carried over from one body to another by connecting them. On the other hand, there are substances that will conduct *electricity*. We are familiar with this fact, for we all know that copper wires are used to conduct electricity to lamps and to other electrical appliances. We know also that these wires are inclosed by materials or mounted on supports which are *nonconductors*, or *insulators*.

When a hard-rubber rod is electrified, only that part which has been rubbed acquires the property of attracting bodies. The electricity does not flow from one part to another, for hard rubber is a nonconductor. But if a metallic body is electrified, the electricity may flow off to the ground. To keep a metallic body from losing its electrified condition it must be provided with supports made of nonconductors. On the other hand, a gasoline truck is commonly kept connected to the ground, usually by a trailing chain. This is to prevent the truck from becoming electrified when gasoline flows from the faucet. All gasoline pumps are kept well connected to the ground; the rubber hose used in a filling station has a metal lining so that the nozzle will not become electrified.

In general, substances cannot definitely be grouped into the two classes conductors and nonconductors, for there are all degrees of conductors—from the metals, which are the best, down to such substances as sealing-wax, amber, hard rubber, and dry glass, which are so poor that they are called nonconductors, or insulators.

373. Charges of electricity. When it was found that the electrified condition could be transferred from one body to another by conductors, it was natural to think of this condition as being due to some kind of fluid on the electrified body, and to conceive the fluid as actually flowing through conductors. This idea gave rise to a number of terms which are still used today. For example, we say that an electrified body has a *charge* of electricity on it and, when it flows from one body to another, that there is an electric *current*. We do not know just what a charge of electricity is, nor do we know what causes the attraction which has been referred to. Whatever it may be, we use the words *charge of electricity*, or simply the word *charge*, to refer to that condition.

374. Fluid theories of electricity. It was long ago suggested that an electric charge was some sort of fluid that could flow over or through conductors. This was a very natural suggestion, for a charge apparently does have some of the properties of a fluid. The *two-fluid* theory assumed that there were two different fluids: one the positive fluid, and the other the negative fluid. These two attracted each other, but either fluid always exerted a repulsive force on the same kind. According to the *single-fluid* theory a charged body was one that had either an excess or a deficit of electric fluid. When a body had the normal amount, it was in a neutral condition. An excess gave it the properties peculiar to one kind of charged condition, while a deficit gave it the properties of the opposite kind.

Both these theories were equally plausible. There was no way to disprove either of them experimentally; one was at liberty to take his choice. The two-fluid theory has been the more popular. There is one important difference between the two theories, which can be made clear by an example. Consider the case of two similar bodies, *A* and *B*, one charged positively, and the other with an equal, but negative, charge. What will happen if the two are connected by a conductor? According to the two-fluid theory either of the fluids might move over to the other body, or both of them might move. The positive might flow from *A* to *B*; arriving at *B*, it would neutralize the negative on *B* and thus leave both bodies neutral. Or the negative of *B* might go over to *A* and accomplish the same thing. Or part of the positive on *A* might go to *B* and part of the negative on *B* go to *A*, the amounts being such as to neutralize both *A* and *B*. According to the single-fluid theory there is only one possible way for the charged condition of both *A* and *B* to be neutralized, and that is for the excess on one to flow over and neutralize the deficit in the other. The condition in this case is more simple and definite.

In order to simplify the descriptions of the process of conduction it became customary to assume that the flow took place only in one direction—from the positive charge to the negative. Thus, in referring to an electric current (for instance, from a dry cell), one says that the current goes from the positive terminal to the negative.

375. The negative-electron and positive-nucleus theory. In section 158 it was stated that it is now known, to a high degree of probability, that each atom of matter is made up of a nucleus with satellites of *electrons*. These electrons are very small, having a mass of about one eighteen-hundredth of that of the hydrogen atom. They are *always* negatively charged, while the nucleus of the atom is *always* positively charged. Some of these electrons are not permanently attached to the atom; they not only can leave the atom but can even leave the body. According to the electron theory a negatively charged body is one which has an excess of electrons; a positively charged body is one in which there is a deficit of electrons,—not enough to neutralize the effects of the positively charged atomic nuclei. When two substances are electrified by rubbing them together, apparently electrons are rubbed off from one on to the other. When hard rubber and wool are used, the wool loses electrons.

When the electron moves, it carries with it its negative charge; in fact, according to the electron theory, this is the only way in which negative charges can move. On the other hand, positive electricity exists only on the nuclei of atoms; hence positive electricity is moved only when atoms are moved. While the electrons in metallic substances are free to move about in the interstices between the atoms and molecules, the atoms cannot move freely. If they did, silver atoms would be carried over into a copper wire, or copper atoms into a silver wire, when silver and copper wires are joined and an electric current is sent through the wires across the junction. There is sometimes found a slight diffusion of one kind of atoms over into the other metal, but no such effect is observed as would be produced if the current were carried by the migration of atoms. The conclusion is drawn that in metallic conductors electricity is carried only by the migration of electrons. Since the electrons carry only negative charges, it follows that in metals only negative charges move. But in the conduction of electricity through certain liquids (electrolytes) it is found that silver or copper may be carried by the current (electroplating). In that case both atoms and electrons move, and hence both positive and negative charges move.

The convention universally adopted long ago, that a current always flows from the positive terminal to the negative, is from the point of view of the electron theory incorrect. Since it is only the negative charge that flows in metals, the movement is from the negative to the positive. However, if we regard the usual statement as only an arbitrary convention, we shall find that its use leads us into no errors.

376. The leaf electroscope. While the detection of a charge can often be made by the ability the charged body has of attracting light bodies, or by the spark produced when the hand is brought near it, yet these methods are not very sensitive. The *leaf electroscope* is a simple, convenient, and very sensitive instrument for detecting the presence of a charge of electricity. There are several different types, but the one shown in Fig. 225 is commonly used. The "box" is made of metal, and has glass ends so that one can see the leaves. The leaves, usually of aluminum or gold foil, are hung from the lower end of a metal rod which passes up through the box and terminates in a knob or a binding-post. A very important feature is that this rod does not make metallic contact with the box, but is separated from it and supported by a collar of amber or other insulating material. This collar prevents any charge which may be on the knob, rod, or leaves from flowing over to the box.

If a charged body is brought into contact with the knob, some of the electricity will be conducted to the knob and thence into the leaves. When the leaves become charged, they will repel each other and swing apart. The instrument is sensitive because the leaves are extremely light.

If the electroscope is not charged, the leaves will diverge when either a positive or a negative charge is brought near. Hence if the leaves initially are not charged, one cannot tell whether the body which is brought up to the knob is positively or negatively charged. If, however, the electroscope is charged so that the leaves are diverging, then the leaves will diverge still more when a body electrified with one kind of electricity is brought near it; while a body having the other kind of electricity will cause the leaves to come together. Hence by using a charged electroscope one can tell what kinds of electricity are on different bodies. The electroscope must first be tested with a known kind of electricity in order to find out whether it is a positive or a negative charge that increases the divergence of the leaves.

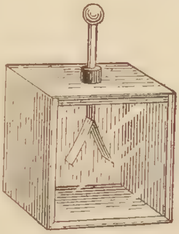


FIG. 225

377. Both kinds of electricity produced simultaneously. It is not difficult to show with an electroscope that when a rubber rod has been rubbed with a woolen cloth both the rod and the cloth are electrified, and that the rod and the cloth have *opposite* kinds of electricity on them. If the rod causes the charged leaves to diverge, the cloth will cause them to converge: the two always have opposite effects. The same thing is true when glass is rubbed with silk. If hard glass is used, it will be found to be charged positively, and the silk negatively.

If a hard-rubber rod and a woolen cloth, both in an unelectrified condition, are rubbed together while near an electroscope, they will show no effect on the electroscope until they are separated. The reason for this is that the two kinds, being produced in *equal* amounts, completely neutralize each other. The conclusions reached, not only from such experiments but from others which are more carefully performed, are (1) that *two kinds of electricity are always produced simultaneously*; (2) that *the amount of positive electricity produced is always equal to the amount of negative electricity produced.*

The student should readily see that these two results *follow from the electron theory.* When the electrons are rubbed off from one substance on to another, the gain in negative charge of one must be equal to the excess positive charge left in the other.

378. Electrostatic induction. In Fig. 226 are shown two metallic conductors, *A* and *B*, each mounted on an insulating support. Let us suppose that *A* had been charged with positive electricity before it was brought near *B* and that *B* had had no charge on it before *A* was brought near, being in what we have called a neutral condition. After *A* is brought near *B*, but not near enough for a spark to jump across nor for any possible conduction of electricity from *A* over to *B*, it will be found that *B* is charged in the manner shown in the figure. On the side nearest *A* there is a negative charge, and on the opposite side a positive one. The simplest way of testing this is to bring a small conductor, mounted on a nonconducting handle, into contact with some point on *B* and then carry it to an electroscope, where the kind of electricity on it can be tested. By trying various points on the surface of *B*, one finds that *B* is charged as shown in the figure. The charges on *B* are said to be produced by *induction.* It should be clearly understood that there has been no *conduction* between *A* and *B*,—that the charges on *B* did not actually come from *A*.

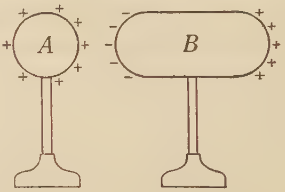


FIG. 226

If *A* had been charged negatively instead of positively, a positive charge would have been found on the side of *B* next to *A*, and a negative charge on the remote side of *B*. In general, *an unlike charge is found on the side next to the charged body, and a like charge on the farther side.*

The foregoing facts can easily be understood and remembered by aid of the electron theory. According to this theory an uncharged body is one with the normal number of electrons (the total negative charge on the electrons being equal to the total positive charge on the atomic nuclei), a positively charged body one with a deficit of electrons, and a negatively charged body one with an excess of electrons. When *A* is brought near *B*, since unlike electricities attract, the positive charge on *A* attracts the electrons in *B*, and, being free to move, they move toward the side of *B* nearest *A*. This leaves a deficit of electrons on the farther side of *B*. In this way one side of *B* acquires a negative charge, and the other side a positive one.*

In this process of charging *B* by induction no charges are added to *B*. Therefore we should expect to find the amount of positive electricity on it still equal to the amount of negative on it. This is shown when, on taking *A* away, we find that *B* is no longer charged: the two kinds on *B*, attracting each other, have mixed and neutralized each other.

However, it is possible to give to the conductor *B*, by induction, an excess of one kind of electricity; that is, a conductor can be charged either positively or negatively by induction. This process is as follows: While *B* is near *A*, and the charges in *B* are separated in the manner shown by Fig. 227, (*a*), *B* is connected to the ground (Fig. 227, (*b*)) by touching it with the finger; and after taking the finger away, *A* and *B* are separated. Then *B* (Fig. 227, (*c*)) will be found to have a charge opposite in sign to that on *A*. The explanation of this is not difficult; the student should try to explain it.

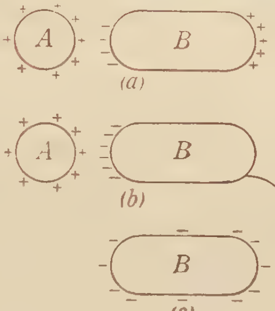


FIG. 227

In this process no electricity goes from *A* to *B* or from *B* to *A*. If it did, would *B* have become charged with the same kind of electricity as *A*, instead of the opposite kind? The air between *A* and *B* is a very good insulator. Air is one of the best insulators (unless the conductors are charged so highly that a spark passes, when the insulation is temporarily broken down). All doubt on the part of the student about the passing of electricity from *A* to *B* or from *B* to *A* can be removed by seeing *B* charged by induction while a glass plate is kept between *A* and *B*.

---

*The student should form the habit of trying to interpret in terms of this simple theory the experiments he sees and those about which he reads. It will help him to remember the facts that he already knows and to predict others. It is important for him to remember that electrons are always negatively charged and that the positively charged particles of a conducting solid cannot move away from their position.

379. An explanation of the action of the electroscope. The action of the electroscope can now be understood. If the electroscope is not charged, and if a negatively charged body is brought near it, the positive and the negative electricity already in it are separated by induction, as shown in Fig. 228. The divergence of the leaves will be produced by the negative charge, which is driven down by the repulsion of the charge on the rod. If a positively charged body is brought near the electroscope, the effect will be that shown in Fig. 229.

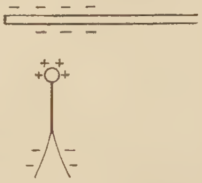


Fig. 228

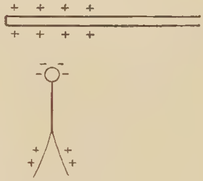


Fig. 229

If, however, the electroscope is charged with positive electricity before a negatively charged rod is brought near, the positive charge in the electroscope will be attracted* by the negative of the rod, and the divergence of the leaves will decrease (Fig. 230). When a positively charged body is brought near the positively charged electroscope (Fig. 231), the positive charge in the electroscope is repelled to the leaves and causes them to diverge more. The student should try to explain what will happen to a negatively charged electroscope when a negative charge is brought near it and also when a positive charge is brought near. (It should now be clear to the student why the leaves must be connected to the knob by a good conductor.)

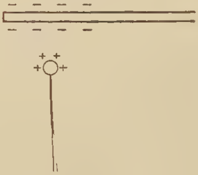


Fig. 230

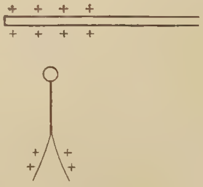


Fig. 231

---

*Conventional language is used here as well as later. The student should be able to interpret this in terms of the electron theory, which states that only negative electricity can move in metals. In this case some of the electrons will be repelled to the leaves and neutralize the positive there.

The metallic case surrounding the leaves of the electroscope plays a very important part in making the instrument more sensitive. A charge on the leaves induces on the inside of the case an opposite charge, which attracts the leaves and increases their divergence.

380. Charging an electroscope by induction. It is often not safe to charge a leaf electroscope by conduction. If the charge happens to be a little too large, the leaves will be separated so violently that they may be torn. Not only is it safer to charge by *induction* but usually it is easier. The process is as follows:

1. A charged body is brought near the electroscope.

2. The knob is touched with the finger, thus removing the charge from the *leaves*. Why?

3. The hand is removed.

4. The charged body is taken away.

What would be the procedure if you wished to give, by induction, a negative charge to the electroscope?

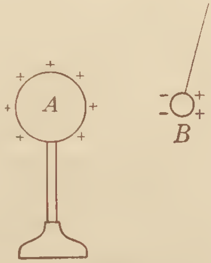


FIG. 232

381. The attraction of pith balls and other bodies explained. When a charged body A is brought near a pith ball B, it acts on it by induction, separating the charges, as shown in Fig. 232. Since the negative charge on B is nearer A than the positive charge, the resultant force is one of attraction. This is the reason why all bodies that are conductors, either good or poor, are attracted by charged bodies.*

382. The return shock. A phenomenon often observed in connection with lightning is due to electrostatic induction. A low-flying cloud charged, say, with positive electricity will induce a negative charge on objects that lie below it—on houses, trees, telephone lines, etc. If this cloud is suddenly discharged to some cloud which was charged with negative electricity, the induced charge on these objects is suddenly released and rushes off to the ground. The tendency of this induced charge to get to the ground quickly causes the crackling around telephones and light-wires during a thunderstorm. The charging process is slow, for the cloud approaches gradually; but the discharge is often very sudden. Sometimes a man in the open may become heavily charged by induction, and receive a shock when the induced charge leaves him. Since this shock is produced by an induced charge returning to the ground, it is called the *return shock.*

---

*It is also true that a small piece of a very good insulator will be attracted by a charged body. The reason for this is not simple, at least not for the elementary student. An explanation is given only in advanced texts.

Just before lightning strikes a tree or house all objects near are charged by induction, and after the stroke they lose their charges very quickly. Many persons who think they have been struck by lightning owe their shock to the induced charge leaving them.

383. The electrophorus. The electrophorus is a simple instrument used as a source of electric charges. There are two essential parts to the instrument: one is a plate of insulating material (hard rubber or sealing-wax) which can be electrified by rubbing with fur or with a woolen cloth, and the other is a flat metal plate which can be brought very close to the charged insulator. Sealing-wax melted and poured into a cake tin serves very well for the insulator. The process of using the electrophorus is as follows: The metal plate, held by a nonconducting handle, is brought near the charged sealing-wax. While there it is touched by the finger; and after the finger has been removed, the plate is taken away. It will then be found charged, often so highly that a spark a centimeter long can be drawn from it. In this process *no charge is taken from the sealing-wax*, and the operation *can be repeated over and over again* without recharging the wax. The action is one of induction, and can readily be understood from Fig. 233. When the conducting plate *A* is brought near the negatively charged wax *B*, positive electricity will appear on the lower side of the plate, and negative electricity on the upper side. Touching the plate will remove the repelled negative charge. After the insulated metal plate has been removed from the neighborhood of the sealing-wax, the positive charge becomes "free" and spreads over the whole surface of the plate.

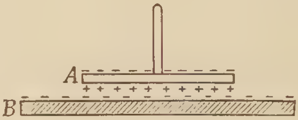


FIG. 233

When the plate is brought near the sealing-wax, it is usually placed directly on the wax, touching it at a few points. The wax is such a poor conductor that an inappreciable amount of its charge is removed. It is because the two can be brought so close together that the inductive action is so strong and that so large a charge is obtained on the plate.

There is a question that is often asked at this point. Some form of it may have already occurred to the student. The conducting plate, when charged, certainly has energy; for the spark produced on the discharge produces some heat and noise. For example, the spark is at a high enough temperature to ignite a gas jet. If this plate does not take any charge away from the charged sealing-wax, where does it get its energy? A similar question might have been asked about the conductor *B* of Fig. 227; for after *B* is discharged, *A* could be brought near again and *B* charged as before, and the process be repeated over and over again without causing *A* to lose any of its charge. The answer is that the energy is supplied by the operator, who does *work* when the plate is removed from the sealing-wax or, in the other case, when *A* and *B* are separated. The negative charge on one attracts the positive charge on the other, and work must be done in separating them. It is a very small amount of work, usually not detected, yet it is enough to give the results described.

Several types of revolving electrostatic machines have been designed which use the principle of the electrophorus. Although producing rather long sparks and attracting much attention when in operation, the electrical power, or rate of developing electrical energy, is usually less than that which can be obtained from an ordinary dry cell. As the complete theory of these machines is rather complicated, the student is referred to larger treatises for a discussion of the details. Since these machines all depend on the principle of charging by induction, or, as it is sometimes called, *influence*, they are often called *induction*, or *influence*, machines.

384. Electric charges from dry cells. The ordinary electroscope is not sensitive enough to detect the charges on a single dry cell. For advanced work most laboratories have electroscopes, such as the *quadrant electrometer*, that are far more sensitive than the common type. In some laboratories one of these may be available for use on the demonstration table. If so, the instructor can show that the two terminals of a dry cell are charged. The central, or carbon, terminal has a positive charge, while the zinc terminal has a negative charge. In the case of a gravity cell the copper terminal is positively charged, and the zinc terminal negatively charged.

If a *battery* is made of seventy-five or one hundred dry cells connected in series (that is, the zinc terminal of one is connected to the carbon of the next, and so on), it is very easy to show with a good leaf electroscope the charges on the terminals of this battery. In performing this experiment the terminal to be tested is metallically connected to the leaves, while the other terminal of the battery should be connected to the case, as shown in a conventional manner in Fig. 234.*

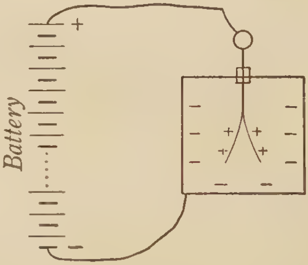


FIG. 234

385. Electric charges from a power circuit. If the electroscope *E* and a suitable lamp *L* are connected to a power circuit (such as the ordinary lighting circuit) by the wires *AB* and *CD*, as shown in Fig. 235, the lamp will be lighted and the leaves of the electroscope will show a slight divergence.

This experiment suggests three noteworthy facts:

1. Although the wires are connected together by the lamp, they are kept charged by the power supply to which they are connected. Touching with the hand does not discharge them.

2. The divergence obtained is small compared with that obtained by using a hard-rubber rod or electrophorus. This is often a great surprise to one testing it for the first time. It will be referred to later, when the whole matter will be cleared up.

3. If the power circuit is a direct-current circuit, the wire connected to what is called the positive terminal is found to have the same kind of electricity as that on hard glass which has been rubbed with silk; and the one connected to the negative terminal has the same charge as that on hard rubber which has been rubbed

---

*In drawings of electrical circuits certain conventional symbols are used. For example, two parallel lines, one long and one short and heavy, represent a cell. A series of parallel lines, as shown in the figure, represents a number of cells connected in series.

with wool. In other words, *the terms "positive" and "negative," as applied to the terminals of cells and power circuits, are obtained from the kind of static charges these terminals carry.*

386. The distribution of charges on conductors. A deep tin can is insulated by standing it on a block of paraffin. While there it can be electrified in a number of ways. A simple way is to introduce into the can a charged hard-rubber rod, care being taken not to touch the can with the rod. While the rod is there, the can is touched with the finger. When the hand is taken away and the rod removed, the can will be found to be charged. If the student cannot readily explain this process and the position of the charges during each step of the process, he should re-read section 378.

After the can is charged, a small metal ball on a silk thread is lowered into it until it touches the bottom. On removing the ball and testing it with an electroscope, no charge will be found on it; but if the ball is touched to the outside of the can, it will become charged. This experiment shows that the charge on the can stays on the outside. This is only an instance of the following general rule: *The charge on a conductor tends to stay on the outside surface.* When we recall that like charges repel each other, we can see why it is that each part of the charge, being repelled by other parts, gets as far from the center of the body as it can.

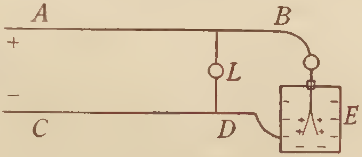


FIG. 235

A sensitive electroscope when placed inside a metal box or cage will not be able to detect the presence of charges that are on the outside of the box. For example, an electroscope will not be affected in any way when inside a wire cage, even when sparks from an electrostatic machine are jumping over to the cage. This can be explained by the fact that the charges on the outside of a conductor do not exert any forces on the inside of that conductor. This is strictly true, however, only when the charges on the outside are at rest. While in motion, they may produce magnetic fields in the inside, as explained in section 405.

As a protection against lightning, part of the efficiency of metal roofs and lightning rods that are well connected to moist ground is due to this tendency of charges to stay on the outside.

If an elongated insulated conductor is charged, and then different parts of the surface are tested with a small conductor mounted on a nonconducting handle, it will be found that there are larger charges on the ends of the conductor. In general, it will be found that those parts of a conducting surface that have the greatest curvature are charged the most highly.

387. The action of points. In the last section it was shown that the charges on conductors have a tendency to accumulate on the more pointed parts. When the end of the conductor is a *sharp* point, the effect of a large charge is so great that the air in the neighborhood of the point becomes a conductor* and then becomes charged with the same kind of electricity as that on the pointed conductor. Fig. 236 shows diagrammatically the pointed conductor positively charged and the positively charged air near it. Since like kinds of electricity repel, the charged air is repelled from the point. In this way the point is very quickly discharged. If the point is kept charged, say by an electrostatic machine, there will be a stream of charged air driven off from the point. This stream of air is often strong enough to blow out a candle flame. Conductors brought into this stream of air become charged; for example, an electroscope fifteen feet away from a point attached to an electrostatic machine can be charged so rapidly that one can see the leaves move.

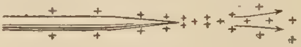


FIG. 236

If an uncharged conductor *B*, with a sharp point on it, is brought near a charged conductor *A*, the conductor *B* will become charged with the same kind of electricity as that on *A*. The reason for this can be understood by the aid of Fig. 237. The point becomes charged by induction with negative electricity, which is conducted away by the stream of air which flows from the charged point toward *A*. This leaves a positive charge on *B*; the negative going to A neutralizes part of the positive charge on A. No matter whether the point is on the charged or uncharged conductor, the effect of the point is to make the air between the two a conductor, and the final result is the same as if the two conductors were brought into contact.

---

* A gas does not conduct in the same manner that a metal does. Air becomes conducting when some of the molecules become charged. Electricity is carried through the gas by the motion of these charged molecules (see section 453, also Chapter XXXVII). If a gas is positively charged, it has more positively charged molecules (ions) than negatively charged ones.

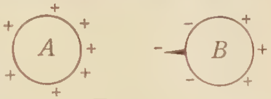


FIG. 237

The *electric whirl* (Fig. 238) depends for its action on the fact that when the electrified air near a point is repelled there must be a *reaction* on the conductor. The whirl consists of a light cross of wire, with ends sharpened and bent at right angles so that they all point in the same sense. When this is mounted on a pivot at its center and kept charged, the repulsion between the charged air and the points causes it to rotate in a direction opposite to that in which the points are bent.

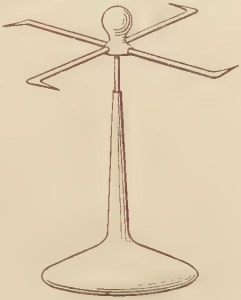


FIG. 238

The fact that points can make the air near them conducting is utilized in many ways. In electrostatic machines rows of metallic points are used to conduct the electricity off from moving plates to stationary conductors. Part of the efficiency of lightning rods is due to the fact that the pointed tips can produce a silent discharge between the rods and low-flying clouds. Sometimes a heavily charged cloud will induce an opposite charge on buildings and other objects lying below it. Sharp points will tend to send this induced charge up into the air toward the cloud and thus to discharge the cloud. However, this action is slow, and hence it is not always effective.

388. The lightning rod. In section 386 it was pointed out that on account of the tendency of charges to stay on the outside of conductors a well-grounded system of lightning rods on a house is of some protection to the interior. In the last section the reason for putting points on the rods was explained. There is, however, considerable difference of opinion about the advantage of installing lightning rods. It is usually a difficult matter to get them well connected to the earth, to wet soil, and it is also difficult to keep them so connected. Electricity in motion does not obey the same laws as when at rest; for example, a charge in motion may prefer to go to the interior of a conductor. We now know that very sudden rushes of charges obey special laws, and that sometimes the discharge will not follow the path that appears to be the best. It may not follow the rod down to the ground, but it may jump across and go down a chimney or may enter the house to some plumbing that comes too near the rod. When installed with reference to these points and to others which are known, a system of rods does afford reasonably good protection. On the other hand, a poorly installed system is of doubtful advantage and may even be dangerous.

389. The law of force between two small electrified spheres. Coulomb was one of the first to make quantitative measurements of the force between two charged bodies. He found that the force between charges on two small spheres varied inversely as the square of the distance between them, and that a measurement of this force enabled him to estimate the size of the charges on the small spheres. In its final form Coulomb's law states that *the force between two small charges is directly proportional to the product of the sizes of the charges and inversely proportional to the square of the distance between them.* This law may be stated in algebraic form as follows:

$$F = k\frac{q_1 q_2}{d^2}, \tag{7}$$

where F is the force, q_1 and q_2 the charges, d the distance between the charges, and k a constant which depends on the units used and on whether the charges are in air or in some other medium, such as oil.

390. Unit charge. Coulomb's law affords a method of measuring charges; but we must first adopt a unit, or standard, charge. The magnitude of the constant k in equation (7) depends on what units are used for measuring F and q.* We shall choose our unit of charge so that k will be equal to unity when the experiment is performed in air. The equation then becomes

$$F = \frac{q_1 q_2}{d^2}, \tag{8}$$

where the force is measured in dynes and the distance in centimeters.

---

* See the remarks on a similar constant in section 65.

It is possible, by a method that has been used before, to define the size of a unit charge by taking a special case of equation (8): *A unit charge placed 1 centimeter away from an equal charge will repel it with a force of 1 dyne.*

This definition and equation (8) are true only when the surrounding medium is air. This will be referred to later.

The unit we have just defined is called the *C.G.S. electrostatic* (C.G.S.E.S.) unit. Another and much larger unit of quantity of electricity, called the *C.G.S. electromagnetic* (C.G.S.E.M.) unit, will be defined later. There is still another, called the *coulomb*, which belongs to the so-called *practical* system of units. The practical units are merely convenient multiples of the electromagnetic units (see section 522).

30,000,000,000 C.G.S.E.S. units = 1 C.G.S.E.M. unit.

10 coulombs = 1 C.G.S.E.M. unit.

391. A natural unit of electric charge. The fact that an electron always has the same-sized charge on it leads to the suggestion that this is the fundamental charge of nature,—the smallest which can exist. This idea has received complete confirmation by the experimental work of R. A. Millikan, who has found that the charge on very small bodies is always, so far as he has tested it, a whole multiple of the electronic charge. The charge on Millikan's small particles never contains any fractional parts of the fundamental charge: the observed charge is always some integral number of electronic charges. Hence it appears that this charge, which may be said to be the *natural* unit of electricity, can never be subdivided. For this reason it is sometimes called an electrical *atom* (a word that means something which cannot be divided). This natural unit is a very small charge. According to Millikan's measurements, which are the most accurate to date, it is equal to 4.774×10^{-10} electrostatic units, or 1.591×10^{-19} coulombs.

The positive charges on the nuclei of atoms are believed to be integral multiples of the natural unit and to vary in magnitude from 1 natural unit on the hydrogen nucleus to 92 on the uranium nucleus.

392. The intensity, or strength, of an electric field. In the region around a charged body there is what is called an *electric field.* If a small charged body (a charged pith ball, for example) is brought to some point in this field, there will be a force acting on the pith ball. An electric field, then, is a region in which forces will act on charged bodies if they are brought into it.

Some method must be devised for defining the *intensity*, or *strength*, of such a field, for it is frequently necessary to compare two different fields. The usual method is to define the strength, or intensity, of a field by saying that *the intensity, or strength, of an electric field is numerically equal to the force that the field would exert on a unit positive charge.*

If the force exerted on a unit charge by a certain field is 10 dynes, then the intensity of the field is said to be 10; if the force on a unit charge placed at another point is 12, then the intensity at that place is 12. In this book the symbol $\mathcal{F}$ will be used to denote the intensity, or strength, of an electric field.

It follows from the definition of the intensity $\mathcal{F}$ of a field that a charge q placed in a field at a point where the intensity is $\mathcal{F}$ will experience a force given by the equation

$$\text{Force on a charge } q = \mathcal{F}q. \qquad (9)$$

393. Electrical lines: the strain theory. It is sometimes convenient, as in the case of magnetic fields, to represent an electric field by lines.

1. A line at any point is always drawn in the direction of the intensity of the field; that is, in the direction of the force which would act on a positive charge if it were brought to that point.
2. The density of the lines at any one place is determined by the intensity of the field at that place.*

In Figs. 239 and 240 the electric field is represented by means of these lines, which always start on a positively charged body and end on a negatively charged one.

If one assumes that these lines have certain properties, the attractions and repulsions of charged conductors can be readily explained. These assumptions are as follows: (1) the lines are under tension and therefore tend to contract; (2) the lines repel each other sidewise.

It has been suggested that the whole phenomenon of charged bodies is, rather, one of the medium: that the medium around the bodies is actually strained in just the way we have imagined these fictitious lines to be. What we have called a charged body is, according to this suggestion, a body with the surrounding medium in a strained condition. This theory is often called the *strain theory* of the electric field, or sometimes the *strained-dielectric theory*. In some advanced treatises it is shown that all the facts of attraction between electric charges can be explained by means of the elastic forces of a strained medium.

---

*The student will notice the similarity between electric lines and magnetic lines. But they must not be confused.

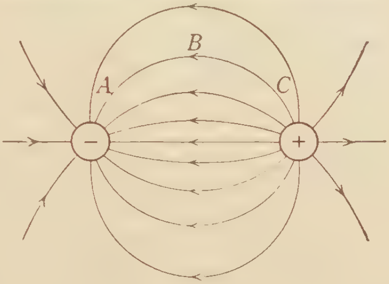


Fig. 239

394. Faraday's ice-pail experiment. The classical ice-pail experiment of Faraday shows that the *total* induced charge due to any charged body is *always equal to the inducing charge*. Or, stated in terms of electrical lines, the lines from a charged conductor reach out until they meet other conductors, and the total charge at the ends of these lines is equal, and opposite in sign, to the charge on the conductor from which the lines start. For example, when an insulated positive charge is brought into a room, there is brought along with it an induced negative charge on the floor, walls, ceiling, and objects in the room. This induced negative charge is equal to the positive charge on the insulated conductor. We shall now see just how Faraday proved this.

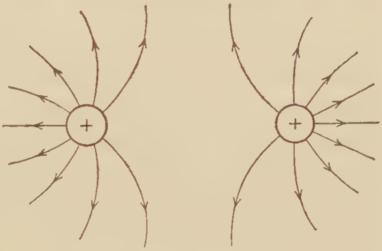


Fig. 240

Faraday's experiment can be readily understood from Figs. 241 and 242. The uncharged insulated conductor A—an ice pail in the original experiment—is connected to an electroscope E by a wire. When a charged ball B is lowered into A by a silk thread, without touching A, charges are induced on A and on the electroscope, as shown in Fig. 241. If the ball is removed, the leaves of the electroscope collapse, showing that A is left uncharged. When the ball is replaced, the leaves of the electroscope diverge as before. If the ball is now touched to the bottom of the vessel A (Fig. 242), the diverging leaves of the electroscope *will show no change whatever.* Moreover, when the ball is removed, *no charge* will be found on it. The explanation is that when the ball touched the bottom of the vessel, the positive on the ball and the induced negative on the inside of the vessel united. The fact that *no charge is left on the ball* and *no change observed in the charges on the outside of the vessel* shows that the two—the positive on the ball and the negative on the inside of the vessel—must have been of the same amount. If they had not been equal, the excess charge would have gone to the outside (for "free" charges always go to the outside of a conductor) and would have increased or decreased the charge on the outside. Any change in the outside charge would have been shown by a change in the divergence of the leaves of the electroscope. Hence the induced charge on the inside of the vessel must have been equal in size to the charge on the ball.

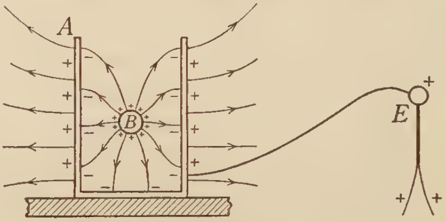


FIG. 241

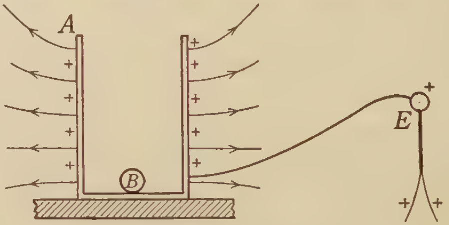


FIG. 242

In the experiment just explained the vessel almost completely surrounds the ball, so that practically all the induced charge is on the inside of A. In the case shown in Fig. 226 (p. 386) the charge induced on B is not equal to that on A. Part of the charges induced by A are on the floor, walls, table, etc. The student should clearly understand that it is the *total* induced charge which is equal to the inducing charge.

395. The Leyden jar. The Leyden jar forms an interesting example of the application of the principle of electrostatic induction.

The inside and outside of a cylindrical glass jar are coated with metallic foil on the bottom and about two thirds of the way up the sides. The inside coating is usually connected by a rod to a knob which projects above the jar (Fig. 243).

If the inside coating of the jar is connected to the positive terminal of some source, such as an electrostatic machine, and the outer coating to the negative terminal, relatively large charges of electricity will be given to the coatings on the jar.

A simple explanation of this is readily deduced from the principle of induction. Since the inner coating is connected to the positive terminal, it will become positively charged. Being positively charged, it will attract negative electricity to the outer coating by induction; and since the outer coating is connected to the negative terminal of the source, the negative charge will flow to the outer coating. The negative charge on the outer coating tends in turn to attract more positive electricity to the inside, and so on, with the result that relatively large charges may be given to the coatings.

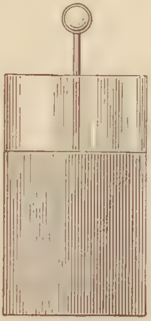


FIG. 243

If the jar is charged by being connected to an electrostatic machine, a heavy discharge, accompanied by considerable noise and light, can be obtained. The jar is discharged by a conductor one end of which is brought into contact with the outer coating, while the other end is brought near the knob.

A more complete explanation of the *condenser*, of which the Leyden jar is one type, is given in Chapter XXXV.

396. A brief history of early electrical discoveries. No one knows when it was discovered that bodies can become electrified and acquire the property of attracting light objects. Certainly Thales, about 600 B.C., knew something about it. Amber, a form of resin, was one of the substances that the ancients knew could be charged. The Greek name for amber is *ἤλεκτρον* (*electron*), and from that word William Gilbert, in 1600, coined the word *electric*. A few years later the noun *electricity* was used by Sir Thomas Browne. Gilbert seems to be the first who endeavored to carry out electrical experiments in what we now call a scientific manner. His attitude doubtless had much influence on his successors.

The first electric machine was made in 1663 by Guericke, the inventor of the air pump. It consisted of a globe of sulfur mounted on an axis. When pressed with the dry palm of the hand while it was being rotated, it became electrified. Guericke found that charges could be carried from one body to another by *contact*, or, as we now say, by *conduction*. A clear distinction was not made between conductors and insulators until 1729, when Stephen Gray showed that a wet string or a wire would conduct electricity. He found also that some substances had insulating properties, and that by means of these substances conductors could be insulated. He transmitted a charge from a rubbed glass through a string some seven or eight hundred feet long, which was suspended by silk threads. Du Fay, a Frenchman,—a soldier by profession, and later a scientific investigator,—showed that all bodies could be electrified; but his most important work was the discovery, in 1734, that there were two kinds of charges. The kind found on glass when rubbed with silk he called *vitreous*, and the other kind *resinous*. Later Benjamin Franklin suggested the names now used, *positive* and *negative*. Du Fay suggested what is known as the two-fluid theory. In a more complete form it was presented by Robert Symmer, in 1759. The rival theory, the single-fluid theory, was suggested by Franklin, in 1765.

The discovery of electrostatic induction (the charging of bodies by induction) is by some credited to Stephen Gray, by others to John Canton and Johann Wilcke (1753). Before that time it had been noticed that bodies at some distance from a charged body became charged, but the complete process was probably not understood until Canton explained it. Electrostatic induction is now a common method of producing charges, as practically all electrostatic machines use this principle. The forerunner of these machines was the *electrophorus*. Wilcke, in 1762, seems to have had the first idea of it; but Volta, in 1775, developed it.

The similarity of the appearance of a spark and its noise to lightning and thunder led many to suspect the identity of their nature. Benjamin Franklin was the first to make a practical suggestion regarding lightning. If lightning were a discharge similar to that from a charged rod, Franklin thought that it should be possible to conduct it by metallic rods. Hence arose the suggestion, about 1750, of the use of lightning rods. Franklin's ideas were received seriously by some; others were inclined to make light of them. However, experimenters in France, Russia, and England showed that his ideas were correct. It was probably Dalibard who, in 1752, by using an iron pole, first conducted lightning to the ground, near Paris. Richmann of St. Petersburg was killed in 1753 by a discharge which had been conducted into his room by an iron rod. It was about this time that Franklin performed his famous experiment with a kite. This kite, equipped with a sharp-pointed wire, became charged when raised into thunder clouds. This charge was conducted down the wet string and was used to produce sparks, to charge a Leyden jar, and to perform similar experiments. To quote from Franklin's own account: "From electric fire thus obtained spirits may be kindled and all the other electric experiments be performed which are usually done by the help of a rubbed glass or tube, and thereby the sameness of the electric matter with that of lightning completely demonstrated." The first lightning rod to be erected on a building in this country was in Philadelphia, in 1760. The first in England was erected in 1762, on the residence of Sir William Watson.

In a letter of 1747 Franklin describes experiments that showed the action of sharp points on conductors — that points could be used either to charge or to discharge bodies. It was because of these experiments that he used a sharp-pointed wire on his kite and that he suggested the use of sharp points on the tips of lightning rods.

Another important discovery was also due to Franklin. In 1775 he found that the interior of conductors has no free charge,— that a charge had a tendency to stay on the surface of a conductor.

The next important step of that time in electrical discovery was in the method of making quantitative measurements of charges. This was first done accurately by Cavendish, in 1771, but for some reason he never published his results. It was left for Coulomb, in 1785, to rediscover the inverse-square law of force upon which are based the definition of unit charge and the methods of measuring charges.

While the Leyden jar and other types of condensers are examples of electrostatic induction, yet it is true that the Leyden jar and its properties were discovered before the principle of induction was understood. One suggestion is that the discovery of the condenser arose in the attempts to diminish the leakage from charged bodies. To do this the charge was surrounded by an insulator. For example, water was put into a glass bottle and then charged. The water on the inside and the hand on the outside formed the two coatings of a condenser. It is related that in 1746 Cunæus, a pupil of Musschenbroek of Leyden, while holding such a charged body in his hand, touched with his free hand a conductor which was connected to the water. As a result Cunæus was shocked by the discharge. This accident aroused great interest throughout Europe and America. There is good reason for believing that this was not the first discovery of the jar. Von Kleist, bishop of Pomerania, is by many given credit for the discovery in 1745, a year earlier than the discovery in Leyden. The suggestion of coating both the inside and the outside of the jar with tin foil seems to have been made later, by Sir William Watson. The general idea seemed to be that electricity was stored in the bottle very much as water could be stored in it, until Franklin showed that the two coatings were charged with opposite kinds of electricity.

The period immediately preceding the American Revolution was one of intense interest in the phenomena of electrostatics throughout both Europe and America. Public lectures giving popular demonstrations of the properties of electric charges were common. Ebenezer Kinnersley, a friend of Franklin, gave a lecture tour throughout the principal cities of the colonies and met with great enthusiasm. Some of the discoveries of that period attracted wide attention, especially that of the Leyden jar. At that time electrostatics was all that was known of electricity; the properties of the electric current—electricity in motion—were wholly unknown. Notwithstanding the great popular interest and the ability of the workers, the scientists of that period did not make any start on the modern practical applications of electricity. It was only after the discovery, in 1799, of another method of producing electric charges, the voltaic cell, that currents of sufficient duration for experimental purposes could be obtained.

It should be clearly understood that none of the discoveries described, nor subsequent ones, were really due to the work of any one man. Detailed study of these discoveries shows that the process is almost a continuous one: little by little, facts are learned, each one leading toward the final crowning achievement. None of the important discoveries of electricity has been the result of pure accident or chance. It has been only by the long, patient work of our forefathers that we now know so much about electricity and have found so many practical uses for it. None of the great fundamental discoveries was made by the so-called *practical* man. Nearly all this work was what we today call *pure* research, and at the time seemed to have no practical significance. The history of the development of electricity gives convincing evidence of the great value of scientific research and the reasons why we should continue to encourage it.

397. Summary. In order that a student may have even a superficial knowledge of electricity it is absolutely necessary that he become familiar with the terms or names used. To be able on demand to give a correct definition of a term is not sufficient; one must learn to *use* the terms in the proper sense. These remarks apply not only to the terms used in this chapter but also to those in all the chapters which introduce different phases of the subject.

1. When two different substances are rubbed together, they acquire the ability to attract pieces of paper, pith, etc. While they possess this property they are said to be *electrified* and to possess a *charge* of electricity.

2. There are two kinds of charges: one, called *positive*, is the kind produced on hard glass rubbed with silk; the other, called *negative*, is the kind produced on a hard-rubber rod or sealing-wax rubbed with wool.

3. Some substances can conduct charges; these are called *conductors*. Substances that do not conduct are called *insulators*. No sharp line can be drawn between conductors and insulators; for some substances conduct so

poorly that under certain circumstances they may be regarded as insulators, although under different conditions they may be regarded as conductors.

4. Both positive and negative charges are produced simultaneously in equal quantities. Whenever one kind disappears, an equal amount of the other kind disappears.

5. Bodies may be charged by induction. The process of charging by induction is one of the most important things discussed in this chapter. An understanding of the action of the leaf electroscope, the electrophorus, the condenser, and many other electrical appliances can be obtained only by first understanding the simple facts of electrostatic induction (see section 378).

6. Charges tend to distribute themselves over the outside of conductors. At sharp corners or points more charge will be found than elsewhere. In the case of highly charged points the air near becomes conducting, and a blast of electrified air is driven off from the point.

7. Charges located outside a hollow conductor exert no forces on charged bodies inside the conductor. This is the so-called *screening* action. Another way of stating this is to say that an electric field cannot be produced on the inside of a hollow conductor by charges on the outside.

8. The *force* between two small charged bodies is directly proportional to the product of the charges on them and inversely proportional to the square of the distance between them. This inverse-square law is the fundamental *quantitative* law of electrostatics, a law on which all quantitative electrostatic measurements are based.

9. Charges are measured by the force they exert on each other. A C.G.S. electrostatic unit charge is defined as one which will exert a force of 1 dyne on an equal charge placed 1 centimeter away (in air).

10. The *total* charge on a body is the (algebraic) sum of the positive and negative charges on that body. This total amount does not change so long as the body is insulated.

11. All charges are multiples of the so-called natural unit, which is the charge found on an electron.

12. The *intensity*, or *strength*, of an electric field at any place is measured by the force the field would exert on a unit positive charge when brought to that place.

13. The properties of an electric field can be visualized by the aid of *electric lines* (see section 393) similar to the lines drawn for magnetic fields.

14. When an insulated body is charged, there always exists an equal but opposite charge on the walls, floor, table, etc. (see Faraday's ice-pail experiment).

15. The Leyden jar is a special type of electrical *condenser*.

16. The negative-electron and positive-nucleus theory is now generally accepted (sect. 375).

PROBLEMS

1. Find the force between two charges, + 80 and − 20, placed 20 cm. apart. What is the force when they are 40 cm. apart?

2. Find the intensity of the electric field at a point halfway between two charges, + 40 and − 20 electrostatic units, placed 20 cm. apart.

3. Find the intensity of the electric field at a point halfway between two charges, + 40 and + 40 electrostatic units, placed 20 cm. apart.

4. The force acting on a charge of + 20 electrostatic units, when brought to a certain place in an electric field, was 40 dynes. What was the intensity of the electric field at that place?

5. Two charges, of + 90 and − 40 electrostatic units, are placed 30 cm. apart. What is the intensity of the electric field (*a*) at a point on the extension of the line joining them, 20 cm. from the negative charge and 50 cm. from the positive? (*b*) at a point on the other side, 20 cm. beyond the positive charge?

6. Two charges, of + 40 and − 40 electrostatic units, are 10 cm. apart. (*a*) What is the intensity of the electric field at a point 10 cm. from each charge? (*b*) If a charge of + 5 electrostatic units is placed at this point, what will be the magnitude of the force acting on it?

7. A charge of + 25 electrostatic units is carried into a room on an insulated conductor. What will be the total induced charge on the walls and objects in the room?

CHAPTER XXVII

SIMPLE PROPERTIES OF ELECTRIC CURRENTS

The voltaic cell, 398. Unit for measuring current, 399. Resistance, 400. The electromotive force of a cell, 401. Effects of electric currents, 402. Heating effects, 403. Chemical, or electrolytic, effects, 404. Magnetic effects, 405. Mechanical force acting on a wire carrying current in a magnetic field, 406.

This chapter contains only a brief description of the more commonly known effects of electric currents (electricity in motion). The purpose is to give the student a general view of some of the facts and a knowledge of some of the terms used in practical work.

398. The voltaic cell. Most of us know that the common dry cell will give an electric *current*. For example, when a small incandescent lamp is connected to a dry cell by means of metallic wires, the existence of a current is shown by the heating of the lamp filament to incandescence. In order to light this lamp by a cell two things are necessary:

1. The lamp must be connected to the cell by good *conductors* of electric currents. Copper wire is commonly used. In ability to conduct electricity silver ranks first, with copper next. Nonmetallic substances are poor conductors and usually will not carry enough current to light the lamp.

2. The electric *circuit* must be *closed*. A metallic conductor must connect one terminal of the cell to one terminal of the lamp, and another conductor must connect the other lamp terminal to the other terminal of the cell. If the circuit is open at any point, even if the gap is only a microscopic one, the current will not flow. There must be a complete path from one terminal of the cell, through the lamp, back to the other cell terminal. The switch for our electric lights, or the push button of our doorbells, is a device by which the circuit is opened or closed at will.

The ordinary dry cell is only one type of *voltaic* cell. There are many other types of cells in more or less common use. The *gravity* cell, which contains two solutions, one of zinc sulfate and the other of blue vitriol (copper sulfate), is often used in telegraph circuits. The name *voltaic* is applied to all these cells, because Volta, in 1799, was the first to construct a cell that would produce an electric current.

399. Unit for measuring current. The current which flows from a voltaic cell or from any other source is usually measured in *amperes*. For example, ordinary incandescent lamps used in house-lighting have from a fourth of an ampere to one ampere of current flowing through them.

400. Resistance. If, instead of copper wires, iron wires of the same size and length are used to connect a lamp to a dry cell, the lamp will not glow so brightly, thus showing that there is *less* current in the circuit. We say that iron wires have a greater *resistance* than copper wires of the same diameter. When a cell is connected through any closed circuit, the current which the cell produces depends on how much electrical resistance there is in the entire circuit. Long wires have greater resistance than short ones of the same diameter; hence a cell cannot give so large a current through long wires as when short ones are used. An analogous case is a water pump connected to a pipe line which runs out and then back to the pump. The pump can keep the water circulating in this pipe, but the amount of *current* which flows through the pipe will depend, in part at least, on the *resistance* offered by the pipe. The longer the pipe or the smaller the diameter, the less the current that will be produced. So it is with the cell: when longer wires, or smaller ones, or those made of a poorer conducting material are used, less current will travel around the circuit.

Resistance is usually measured in *ohms*.

401. The electromotive force of a cell. The current from a cell or battery depends on two things: (1) the *resistance* of the circuit and (2) the *electromotive force* (or *E.M.F.*) of the cell or battery. The electromotive force is commonly measured in *volts*; for example, that of a dry cell is about 1.5 volts. When two dry cells are connected in *series* (that is, the carbon of one connected to the zinc of the other), the electromotive force of the combination is the sum of the two, or about 3 volts. Such a group of cells is called a battery. Batteries for flash lights usually consist of two or three dry cells connected in series.

The current in a circuit is directly proportional to the electromotive force and inversely proportional to the resistance of the circuit. Or the number of *amperes* is directly proportional to the number of *volts* and inversely proportional to the number of *ohms* of the complete circuit. This rule, commonly called *Ohm's law*, will be discussed more fully in a later chapter.

402. Effects of electric currents. The more commonly known effects of electric currents can be grouped under the following heads:

1. Heating (and, as a result, lighting) effects.
2. Chemical or electrolytic effects.
3. Magnetic effects (as in electric bells and telegraph sounders).
4. Mechanical force acting on wires carrying currents in them (as in an electric fan or in the starter of an automobile which is run by such force).

These will be explained briefly, in order. The student should note carefully that these effects are due to what is called an electric *current*, to *electricity in motion.*

403. Heating effects. The heating of a lamp filament by the current flowing through it has already been mentioned. In the case of an electric iron or an electric stove the current flowing through coils of wire develops heat in them, often making them red-hot.

The precise laws which state the amount of heat developed will be given later. One point should be noted here: for the same amount of current a good conductor is heated less than a poor one. When the same current flows through a copper wire and an iron wire of the same size, greater heat is developed in the iron wire. It is for this reason that the heating-coils of electric irons, stoves, etc. are made of relatively poor conductors. For the same current the rate of development of heat is proportional to the resistance of the conductor.

404. Chemical, or electrolytic, effects. If two plates of copper are placed in a solution of blue vitriol (copper sulfate) and connected to several dry cells so that a current flows through the solution from one plate to the other, copper will be deposited on one plate and dissolved from the other. If, instead of the copper plate on which the deposit forms, a plate of some other metal is used, copper will be deposited on it. This process is known as electroplating. Gold, silver, and nickel are also readily deposited by an electric current when the proper liquid is used. All the common nickel-plated objects are plated by means of an electric current. Gold-plating and silver-plating, in nearly all cases, are done by the electrical process.

The chemical action of currents was one of the first effects of an electric current to be discovered. In March, 1800, Volta sent a letter to England in which he described the voltaic cell. Within a few weeks Sir Anthony Carlisle and William Nicholson found that water could be decomposed into hydrogen and oxygen. Soon after that, Ritter, of Silesia, found that when a current passed through a solution of blue vitriol one of the wires dipping into the solution was copper-plated.

The subject of electrolytic effects is so important that an entire chapter (XXXI) will be devoted to it.

405. Magnetic effects. Although something was known about magnetism and electricity more than a thousand years ago, no relation or connection between them was discovered until the nineteenth century. There are two fundamental relationships between magnetism and electricity, both of tremendous importance. One of these will be described here; the student should be on the lookout for the other.

In 1819 the Danish physicist Hans Christian Oersted found that when a wire carrying a current was held over a compass needle, the *needle was deflected*, tending to set itself at right angles to the wire. This now ranks as one of the two or three most important discoveries in electricity. One way of describing this discovery is to say that Oersted found that a current flowing in a wire *produced a magnetic field*; for he found that when a magnet (compass needle) was brought near a wire, forces acted on the poles of the magnet, and this is the test for the existence of a magnetic field. The reason that this discovery ranks so high is that most of the applications of electric currents utilize the fact that a current produces a magnetic field. The electric bell, the telegraph sounder, the telephone, all dynamos, and all electric motors depend on the magnetic action of electric currents.

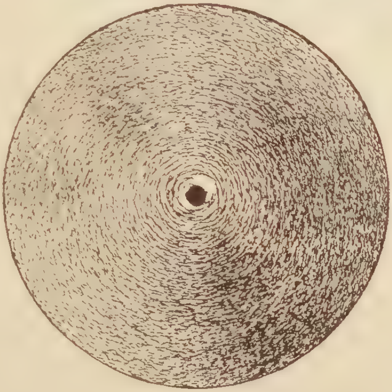


Fig. 244

The best way of representing a magnetic field is by means of magnetic lines. In the case of a current in a wire these are endless lines surrounding the wire. Fig. 244 is a photograph of the lines as traced out by iron filings, the wire being perpendicular to the plane of the paper, passing through at the center of the figure.

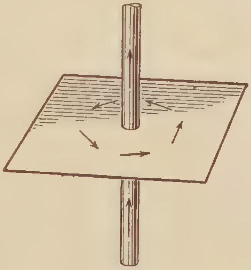


Fig. 245

In section 347 the direction of a magnetic field was stated to be the direction in which the north-seeking pole of a compass needle points. Using a compass needle in the field produced by a current, we find that when the current flows in one direction, the needle points one way around the wire; when the current is reversed, the direction of the needle is reversed. In Fig. 245 the direction of the needle is shown in several positions when the current is flowing up* the wire. In all cases the needle points in the same direction as the magnetic lines around the wire.

A very simple rule for the direction of the magnetic field around the wire is the following: *When a person, actually or in imagination, grasps a wire with his right hand, with the thumb lying along the wire and*

---

* By general agreement the current from a dry cell is said to flow away from the positively charged carbon terminal, which is usually the one in the center, and back to the cell at the negatively charged zinc, or outer, terminal. In the case of a gravity cell it flows away from the copper and back to the cell at the zinc electrode.

pointing in the direction of the current, the fingers will point around the wire in the same direction as the magnetic field.

This rule is very useful in many instances. For example, one can tell by a compass the direction of the field around the wire, for the compass will point in the direction of the field. Knowing, then, the direction of the field, one can at once, by the right hand, determine the direction of the current.

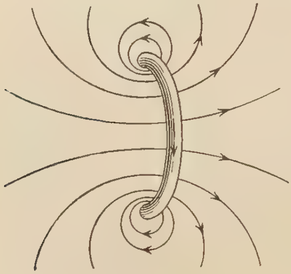


FIG. 246

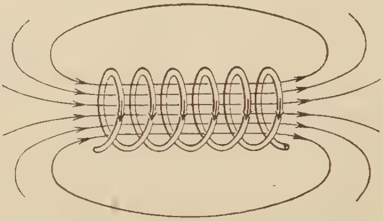


FIG. 247

Figs. 246 and 247 show the character of the magnetic lines produced by a current in a circular coil and in a helix. The direction of the current is marked in each case. By means of the rule just given, the student should verify the directions of the fields shown in the figures.

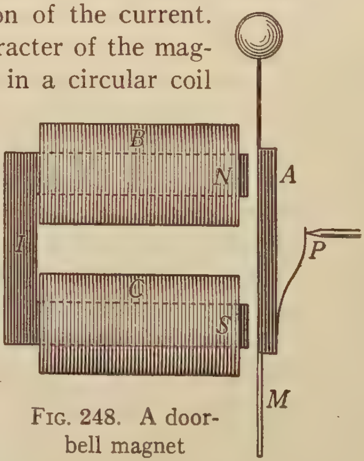


FIG. 248. A door-bell magnet

If a bar of iron were inserted through the coil of Fig. 246 or the helix of Fig. 247 when a current is flowing through the wires, the iron would become magnetized by induction (see magnetic induction, sect. 354). The field of the iron magnet would be added to the field of the coil, with the result that a very strong magnetic field would be obtained. A magnet produced in this way is called an *electromagnet*. If the core is a short piece of soft iron, it will retain its magnetism only while the current is flowing.

Fig. 248 shows a sketch of the magnet of a common doorbell. There are two connected coils, or spools, of wire, *B* and *C*, with an iron core *I*. The current goes around these connected coils in such a direction as to produce two opposite poles, one north-seeking and the other south-seeking. Whenever the circuit is closed, this electromagnet attracts the soft-iron armature *A*, which is supported by the spring *M*. When the armature moves up to the magnet, it opens a contact at *P*, through which the current flows, thus breaking the circuit. As soon as the circuit is broken the magnet loses its magnetism, and the spring *M* pulls the armature back. Then the circuit is closed again and the operation repeated.

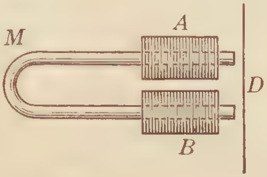


FIG. 249. A telephone-receiver magnet

Fig. 249 shows in section a type of magnet used in a telephone receiver. Two coils, *A* and *B*, are mounted on a magnet, *M*, near its poles. When a varying current flows through the coils, the attraction on the steel diaphragm *D* will change as the intensity of the current changes. The diaphragm will thus vibrate and give out sound-waves which will have the same frequency as the variations in the electric current. Fig. 250 shows a vertical and a horizontal section of a lifting-magnet, such as is used in lifting masses of iron or steel in yards and factories. The frame is of iron with coils of wire *CC* wound around the central core.

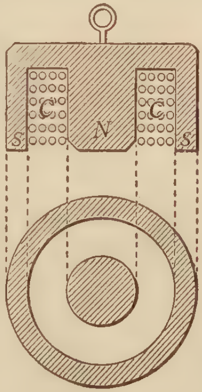


FIG. 250. A lifting-magnet

406. Mechanical force acting on a wire carrying current in a magnetic field. If a coil of wire is suspended near the pole of a magnet, and a current is sent through it, there will be forces acting on the coil which tend to rotate it; if it hangs as shown in Fig. 251, it will tend to turn until it faces the pole. There is therefore a mechanical force acting on a coil of wire through which a current flows when a magnet is brought near. This is also easily shown by the simple experiment of the floating cell. A small test tube containing acidulated water is pushed through a large flat cork. A coil about three inches in diameter and consisting of a hundred or more turns of fine copper wire is connected at one end to a strip of zinc, and at the other end to a strip of copper. The coil *A* and the strips *C* and *Z* are bound with tape or twine so that the coil will hold the shape shown in Fig. 252. When the strips of zinc and copper are dipped down into the acidulated water, a current will flow around the coil. If the cell and coil float on water in a dish and if a pole of a magnet is brought near the coil, the action of the mechanical forces is plainly seen.

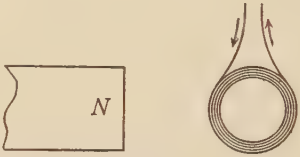


FIG. 251

The tendency of a coil to rotate when placed in a magnetic field has many applications. The ordinary galvanometer, which is used in the laboratory to measure or detect currents, has a coil of wire (Fig. 253) hung between the two poles of a magnet. This coil is suspended by a fine wire which serves to carry the current, the circuit being completed by another fine wire leading from the bottom of the coil. When a current flows around the coil, the latter tends to turn so that its plane will stand perpendicular to the magnetic field

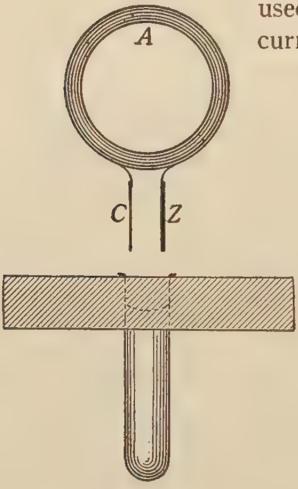


FIG. 252

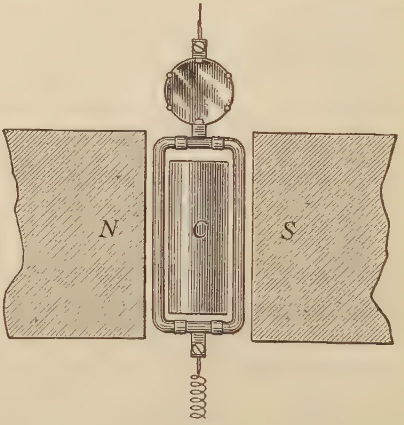


FIG. 253. The moving-coil galvanometer

between the poles; if the current is reversed in direction, the coil will turn in the opposite direction. The mechanical construction of galvanometers has been so highly developed that they are now extremely sensitive; that is, a perceptible rotation will be produced by a very small current.

The *ammeter* (the name of which is an abbreviation of *amperemeter*) is, as the term implies, an instrument used for measuring the current in amperes. The ordinary types of direct-current instruments contain a movable coil hung between the poles of a magnet. In principle the ammeter is like the galvanometer just described. Many electric motors work in a similar way: the *armature*, which contains coils of wire, is placed in a strong magnetic field, and rotates when a current flows through it.

The student should not be content with the meager information contained in this chapter about the mechanical effects of currents. It should be obvious that this is a very important part of the field of applied electricity. The topic is discussed more fully in Chapter XXXII.

CHAPTER XXVIII

POTENTIAL DIFFERENCES; WORK AND POWER IN CIRCUITS

The quantitative relation between quantity and current, 407. Potential: an analogy, 408. Differences of potential: voltage, 409. Voltmeters, 410. Voltages of cells, 411. Resistance; Ohm's law, 412. Resistances of different materials, 413. The voltage drop along wires, 414. Summary of facts regarding the difference of potential, 415. Quantitative definition of difference of potential, 416. Measurement of difference of potential, 417. Work done in a part of a circuit, 418. Power expended in a part of a circuit, 419. Other units for measuring work, 420. Practical examples, 421. Work done in a complete circuit: electromotive force, 422. Summary, 423.

The student should now be ready to learn the *quantitative* meaning of some of the more common concepts in use in practical applications of electricity. In this chapter special attention is given to those involved in the measurement of *power* and *work*, which play so important a part in electrical practice.

407. The quantitative relation between quantity and current. Popularly the word *current*, as applied to the flow of the water in a river, is used in two different senses: sometimes it refers to the *speed* of the water, and sometimes to the total quantity of water passing in a given time. For example, if more water flows during a given time in a wide, sluggish stream than in one that is narrow but swift, the current of the wide stream is said to be larger. With this meaning, the current of the river may be measured in gallons per minute or in cubic feet per minute. In electricity the term is used in this latter sense, referring not to the velocity of the charge but to the total *quantity passing in a unit of time.*

Hence *current of electricity is the rate of flow of charge* or, in the case of a uniform flow, *the charge passing in 1 second.*

If I is the average current, the total quantity of electricity, Q, passing along a wire in t seconds, is given by the equation

$$Q = It. \tag{10}$$

In using this equation both Q and I must be measured in the same system of units. The current I is usually measured either in *C.G.S. electromagnetic* units or in *amperes*. When referring to static charges, we usually measure Q in *C.G.S. electrostatic* units, but rarely when dealing with the flow of electricity. Expressed in practical units equation (10) becomes

$$\text{Coulombs} = \text{amperes} \times \text{seconds}. \tag{11}$$

408. Potential: an analogy. Most of the common applications of electricity today are cases of the transfer or transformation of *energy*. When one turns on an electric light, he connects the lamp to a source of energy. This energy flows into the filament and is there converted into other forms of energy—heat and light. The electric-light meters found in so many houses are instruments that indicate the amount of energy which has been delivered; and it is energy that is paid for, since the unit commonly used, the *kilowatt-hour*, is a unit of energy. The electric stove and the electric motor are instruments for the transformation of electrical energy into other forms of energy. So common is this transformation that it is often necessary to estimate or to compute the amount of electrical energy that is needed. In no other branch of human knowledge are measurements more important than in electricity. But one cannot make estimates or computations of electrical energy without knowing something about *potential*. For this an analogy will be helpful.

Suppose one wished to find out something about the amount of energy available from a given water supply; for example, a reservoir located at some higher place. Could the available energy be determined if one knew only the quantity of water that could be used? No; more information than this is needed. In addition to knowing the quantity of water, one must know the elevation of the reservoir, or the pressure in the water mains due to the height of the reservoir. If the pressure is small, a large quantity of water may be required to supply a certain amount of energy; if the pressure is high, a smaller quantity of water suffices. Two things must be known: the *quantity* and the *pressure*.

A similar condition exists in estimating electrical work. The available energy cannot be estimated merely from a knowledge of the *quantity* of electricity. Something else must be known, something analogous to pressure in the case of water power. This needed concept is called *potential*. In electrical usage the term *potential*, as will be pointed out later, has a definite quantitative meaning; but at present the student may regard it as a term analogous to *water pressure*. A simple case will illustrate. Water always tends to flow from higher to lower elevations or from higher to lower pressures. In a similar way a *positive* charge of electricity tends to flow from higher to lower potentials. In order that one may obtain water from the water mains, the pressure in the mains is kept greater than atmospheric pressure. For a somewhat similar reason the two wires of a house-lighting circuit are kept at different potentials, the difference usually being a little over 100 volts. When the two wires are connected through a lamp, a current flows from the wire at higher potential to the one at lower potential and lights the lamp.

Potential is such an important and fundamental concept in electricity that the student should form a definite idea of its character and use. For this reason he should study very carefully the following sections.

409. Differences of potential: voltage. In considering cases of the flow of water from one elevation to another we are not especially concerned with the elevation of these two places above the mean sea level. What we are interested in is the *difference* of elevation of the two places. The same thing is true of the use of pressures. For example, in stating the pressure of steam or water it is customary to give only the difference in pressure between the steam or water and the air outside. The same sort of thing is true in electricity. In no practical problem do we need to know the absolute potential of any conductor: it is only the difference in potential between that conductor and some other one that must be known.

In practical work, *difference* in potential is usually measured in *volts*, and the difference of potential is often called the *voltage*,— a term that refers to a difference in condition of two conductors. For example, when it is said that the voltage of the light circuit is 110 volts, what is meant is that the difference in potential of the two wires which enter the house is 110 volts.

Before undertaking to learn the precise definition of difference of potential and how it is used in computing work and power, it will be to the student's advantage to learn a number of facts about potential differences. These will be given in the following sections.

410. Voltmeters. Several different types of commercial instruments are now available for measuring difference of potential. The divergence of the leaves of the ordinary electroscope is a measure of the difference of potential existing between the leaves and the case of the instrument. Similarly, the electrostatic voltmeter, in which a delicately balanced charged vane moves when attracted or repelled by another charged conductor, measures difference of potential (Fig. 254). The scales in the commercial forms are ordinarily graduated so that they read *volts*, the indicated number being the difference in potential between two terminals mounted on the instrument. The common type of voltmeter that we find on switchboards and in laboratories also has two terminals, and the deflection of the index indicates the difference of potential between these two terminals.

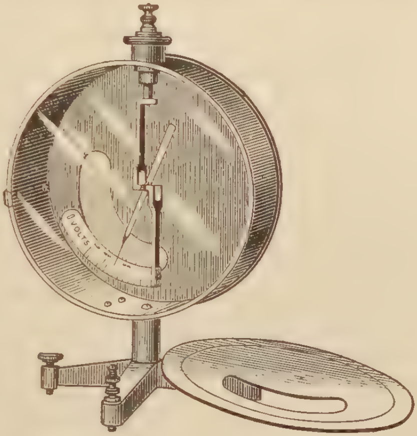


FIG. 254. The electrostatic voltmeter

If for the time being we regard the difference of potential (voltage) as that which is indicated by some form of voltmeter, we can readily ascertain a number of facts regarding potential. If the student wants to find out something about water pressure, he can use a pressure gauge and find the facts by actual test. In a similar way he can use an instrument that indicates potential difference, to test out a number of cases.

411. Voltages of cells. In section 384 it was pointed out that the terminals of a dry cell are charged. The experiment described there showed also that the terminals are not at the same potential; for the instrument used, the electroscope, indicates difference in potential. But it is much easier to show that the terminals of the cell are not at the same potential by using an ordinary voltmeter.

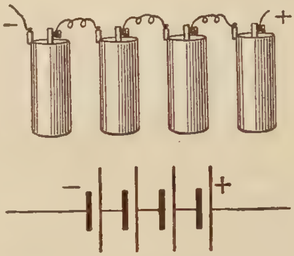


FIG. 255

1. The difference of potential between the two terminals of a dry cell is about 1.5 volts. Touching these terminals does not discharge them. If a small lamp is lighted by connecting it across them, the terminals will still be at different potentials, since these potentials are maintained by chemical action in the cell.

2. In section 384 an experiment was described which showed that the carbon terminal of a dry cell was positively charged, and the zinc terminal negatively charged. Since positive electricity tends to go from a higher to a lower potential, it follows that the carbon terminal is at a higher potential than the zinc terminal.

3. When a number of dry cells are connected in series, larger differences of potential are obtained. To connect them in series the carbon terminal (usually the central one) of the first cell is connected to the zinc of the second, the carbon of the second to the zinc of the third, and so on. A voltmeter will show that if each cell has a difference of 1.5 volts, two cells will have twice that, three cells three times, and so on. The student should have very little difficulty in seeing why this is true. The carbon of the first cell in Fig. 255* is 1.5 volts higher than the zinc terminal, and the carbon of the second cell is 1.5 volts higher than the second zinc. But the zinc of the second is at the same potential as the carbon of the first; hence the carbon of the second is 3.0 volts

---

*The figure shows also the conventional method of drawing cells.

higher than the zinc of the first. Batteries containing large numbers of cells have been built. With 10,000 dry cells in series a difference of potential of about 15,000 volts is obtained. This relatively large difference in potential will give a spark discharge about half a centimeter long. At such high potentials the terminals attract pith balls and show other properties of static charges.

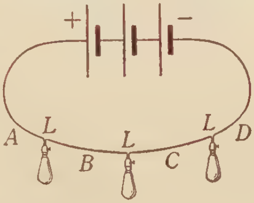


FIG. 256

4. In Fig. 256 is shown diagrammatically a battery of three cells connected to several incandescent lamps, *L*, arranged in series. If one connects up a circuit in this way and then, with a voltmeter, tests the voltage across *AD*, *BD*, and *CD*, he will find that the potential of *A* is higher than that at *B*, that at *B*, in turn, higher than that at *C*, and so on. In other words, there is a progressive *fall* in potential from *A* through *B* and *C* to *D*. The current from the positive terminal of the cell, in flowing from *A* to *D*, is flowing to places of lower and lower potential. Imagine that one could start at the positive terminal of the battery and travel around this circuit from *A* to *B*, to *C*, to *D*, and on through the cells, and that he could in some way measure the potentials at various points. He would find that the potential steadily fell as he passed from the positive terminal around through *A*, *B*, *C*, and *D* to the negative terminal, but that in the cells, where the chemical action was taking place, the potentials increased. An analogous case is of interest. Fig. 257 represents a water pump, *P*, forcing water around a closed circuit of pipe. Suppose that the pipes are all at the same level, so that gravitational forces do not enter. There is a continuous fall of pressure from *A* around to *D*. Inside the pump the pressure is again raised. There is a loss of pressure from *A* around to *D* only when there is a flow, or current. The corresponding fact in an electric circuit is very important—there is a fall, or loss, of potential in an electric circuit only when a current is flowing.

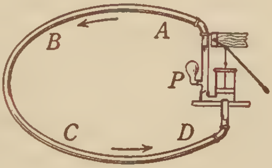


FIG. 257

Potential and the fall of potential are analogous to pressure and the fall of pressure in flowing water. But the student must not conclude that potential is the same thing as pressure, any more than he should decide that electricity has all the properties of water.

412. Resistance; Ohm's law. The resistance of a wire or any part of an electric circuit is so defined and the units for measuring it are so chosen that the following relation holds: *Fall of potential in a part of a circuit is equal to the current times the resistance of that part.* Expressed in symbols,

$$V = IR. \tag{12}$$

When V, the fall of potential, is measured in *volts*, and I, the current, in *amperes*, the resistance R is measured in *ohms*.

Resistance of any part of a circuit is therefore defined as the factor which, when multiplied by the current, gives the loss of potential in that part of the circuit. The numerical value of the ohm is so chosen that the product of the resistance in ohms and the current in amperes is equal to the fall of potential in volts.

It is found by very careful experiments that in nearly all cases *resistance does not depend on the magnitude of the current*, provided the current is not large enough to produce physical changes in the material, such as a change in temperature. If a coil of wire has a resistance of 10 ohms for a current of 5 amperes, it will have the same resistance when the current is only a small fraction of an ampere.

Equation (12) is often called *Ohm's law*. It may be written in terms of the names of the units in any of the following equivalent forms:

$$\text{Volts} = \text{amperes} \times \text{ohms}, \tag{13}$$

$$\text{Amperes} = \frac{\text{volts}}{\text{ohms}}, \tag{14}$$

$$\text{Ohms} = \frac{\text{volts}}{\text{amperes}}, \tag{15}$$

where the difference in potential of *part* of a circuit is measured in volts, and the number of ohms is the resistance of that part.

The second of these forms, equation (14), appeals to many as being the easiest to interpret. This equation states that the current is directly proportional to the voltage of a part of the circuit and inversely proportional to the resistance of that part.

Examples. 1. An electric iron has a resistance of 20 ohms. What current will flow through it when it is connected to a light-circuit which has a difference of potential of 110 volts? (From equation (14) the current is 110/20, or 5.5 amperes.)

2. An incandescent lamp has 0.5 ampere when connected to wires which have a voltage of 110 volts. What is the resistance of the lamp? *Ans.* 220 ohms.

413. Resistances of different materials. An iron wire and an equal length of copper wire of the *same diameter* are connected in series so that the same current flows through both wires. If, by means of a voltmeter, V (Fig. 258), one measures the difference of potential between the two ends of the iron wire and then, either with the same voltmeter or a different one, measures the difference of potential between the ends of the copper wire, he will find that the fall of potential along the iron wire is considerably larger than that along the copper wire. For the iron wire

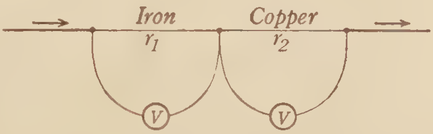


FIG. 258

$$\text{Fall of potential} = I \times r_1,$$

where r_1 is the resistance of the iron wire; for the copper wire

$$\text{Fall of potential} = I \times r_2,$$

where r_2 is the resistance of the copper wire, and I is the current (which is the same in both wires). The fact that the fall of potential is greater in the case of the iron shows that the resistance of iron wire is greater than that of copper wire.

The following table gives the *relative* resistances of different materials for the same length and cross section:

Silver	1.00	Tin	7.
Copper	1.08	Platinum	7.
Gold	1.4	Soft steel	10.
Aluminum	1.8	Lead	13.5
Tungsten	3.	Mercury	60.

The numbers in this table should be regarded as only approximate, for the resistance of a wire will usually depend not only on the purity of the metal but also on its condition; for example, on whether it is hard-drawn or annealed.

414. The voltage drop along wires. Copper wire, No. 12 American wire gauge, about 1/12 inch in diameter (a size suitable for house-wiring for electric lights), has about 1.6 ohms for each 1000 feet. The *loss in voltage* along such a wire when a current flows through it is given by one form of Ohm's law,

$$\text{Loss in voltage} = Ir, \tag{16}$$

where r is the resistance of the part under consideration (this is sometimes called the Ir loss).

As an example suppose a given house to be connected to the power circuit in the street or alley by 200 feet of No. 12 wire. The resistance of this 200 feet of wire would be about 0.3 ohm. If an electric iron taking 6 amperes were being used in the house, there would be a loss, or drop, in potential along the line of

$$6 \times 0.3 = 1.8 \text{ volts.}$$

If the voltage at the street or alley were 110 volts, it would be $110 - 1.8$, or 108.2, volts in the house. If lamps were also burning in the house, so that the current in the circuit was larger than 6 amperes, then the loss of voltage would be still greater. Can you now explain why turning on an electric iron or a toaster often causes the lights to dim a little?

415. Summary of facts regarding the difference of potential.

1. Positive electricity tends to flow from a point of higher to one of lower potential.

2. Negative electricity tends to flow from a place of lower to one of higher potential.

3. To maintain a flow of electricity through any conductor, there must be a difference of potential between the ends of the conductor, and by some means the difference of potential must be maintained.

4. The "direction of the current" (a conventional expression) is always from higher to lower potential, in the direction a positive charge tends to move.

5. Difference in potential, or voltage, between any two places is usually measured in terms of volts.

6. Electroscopes and voltmeters are used to measure differences of potential.

7. Electrostatic machines, hard-rubber rods, glass rods, etc. give high potentials but cannot maintain them. On the other hand, cells (for example, dry cells) produce small differences of potential but have the ability to maintain them.

8. The difference of potential of the terminals of a dry cell is about 1.5 volts. When cells are connected in series, the voltage of the battery is the sum of the voltages of the individual cells.

9. The fall of potential through any part of an electric circuit is proportional to the current: the greater the current, the greater the fall. It is also proportional to a quantity called the *resistance*. Ohm's law gives the following numerical expression for the relation between the three quantities:

$$\text{Voltage} = \text{current} \times \text{resistance}.$$

10. In a closed electric circuit there is a fall in potential from the positive terminal around through the external circuit to the negative terminal. Inside the cells, where the chemical action is taking place, there are increases in the potential. Outside the cells the current flows from higher to lower potentials, but inside the cells it flows from lower to higher potentials. The fall in potential, or loss of voltage, in any part of a circuit on account of resistance is equal to Ir.

11. Different materials have different electrical resistances. For example, an iron wire has a greater resistance than a copper wire of the same diameter and length. To produce the same current a larger difference of potential is necessary for the iron wire.

416. Quantitative definition of difference of potential. Water flows from higher to lower levels. Positive electricity flows from higher potentials to lower potentials. In each of these cases, when the flow takes place there are changes in energy. If water flows from one reservoir to another at a lower level, the water loses energy, and useful work may be done; or, if the water is raised from a lower to a higher level, work must be done on it. In an analogous way, if positive electricity is carried from a conductor of lower to one of higher potential, work must be done. The expressions for the work in each of these cases may be stated as follows:

1. Work done, or the energy which must be supplied, in carrying water from a level l_1 to a higher level l_2 is given by the equation

$$\text{Energy} = (\text{weight of water}) \times (l_2 - l_1). \qquad (17)$$

2. The energy supplied in carrying positive electricity from a conductor maintained at a potential V_1 to one at a higher potential, V_2, is as follows:

$$\text{Energy} = (\text{quantity of electricity}) \times (V_2 - V_1). \qquad (18)$$

If the charge goes from higher to lower potential, it loses energy and can do work. In either case

$$\text{Work done} = \text{quantity} \times \text{difference of potential}. \qquad (18\,a)$$

Equation (17) might be used to measure difference of levels. If one could measure the work done (that is, the energy supplied in carrying a certain weight from one level to another), the difference in level could be computed from equation (17); for, by this equation,

$$l_2 - l_1 = \frac{\text{energy supplied}}{\text{weight}}.$$

This is not a convenient way of measuring differences of levels. But in the case of electricity the analogous method is one of the best we have for determining the difference of potential. It is the *standard method of defining quantitatively the difference of potential*, or $V_2 - V_1$. From equation (18) we have

$$V_2 - V_1 = \frac{\text{energy supplied}}{\text{quantity}}. \qquad (19)$$

Or, stated in words, *the difference in potential of two conductors is numerically equal to the work done in carrying a charge from one conductor to the other, divided by the charge.* If the charge is a *unit* charge, then we may say that the work done is numerically equal to the difference of potential. Hence *the work done in carrying a unit charge from one place to another is equal to the difference in potential between the two places.*

In the practical system of electrical units the difference of potential is measured in *volts*, the quantity in *coulombs*, and the work in *joules* (1 joule $= 10^7$ ergs). These units may be used with equations (18) and (19).

When 1 coulomb of electricity flows through a lamp connected to a 110-volt circuit, the amount of energy supplied is equal to 110 joules (equation (18)). In this case (that is, when *unit* quantity of electricity is used) the energy supplied is numerically equal to the voltage, or difference of potential.

417. Measurement of difference of potential. All methods of measuring difference of potential must be based directly or indirectly on equation (19), for that equation is the definition of difference of potential. Those methods based directly on this equation are called *absolute determinations* of the potential differences. However, so many absolute determinations have been made that there is now on the market a large number of instruments which have been calibrated; that is, the value of the potential difference, or voltage, that will produce a definite deflection is known, or they are so graduated that the value of the potential differences can be read off directly in *volts*.

For measuring very high voltages special instruments are needed. One method—an approximate but simple and convenient one—is to measure the length of spark which the given voltage will produce. Tables are available giving the value of the voltage which will produce a spark of any given length. Too much reliance must not be placed on the numerical values of the voltages given in such tables: the values depend on the character of the spark gap and on other factors which are not always easily controlled.

SPARKING POTENTIALS

Spark Gap in Centimeters	Between Needle Points, Alternating Voltage. Volts (Maximum)	Between Balls 1 Cm. in Diameter, Direct Current. Volts
0.1	1,000	5,000
0.3	4,000	11,000
0.6	7,000	20,000
1.0	12,000	27,000
2.0	24,000	36,000
4.0	42,000	45,000

418. Work done in a part of a circuit. As stated in section 416, the work done when a charge Q flows from one conductor to another is given by equation (18 *a*)

$$\text{Work} = Q \times \text{difference in potential.}$$

In the case where the flow is uniform the current I is related to the charge Q by equation (10),

$$Q = It,$$

where t is the time required for the quantity Q to pass a point on the circuit. If this value of Q is substituted in (18 *a*),

$$\text{Work} = It \times \text{voltage}, \tag{20}$$

or, stated in practical units,

$$\text{Work in joules} = \text{amperes} \times \text{seconds} \times \text{volts}. \tag{20 a}$$

In a certain electric toaster the current was 5 amperes when the voltage was 105 volts. In 10 minutes, or 600 seconds, the work done was

$$5 \times 105 \times 600 = 315{,}000 \text{ joules.}$$

Since 1 gram-calorie = 4.19 joules, this toaster will receive 315,000 divided by 4.19, or 75,200, gram-calories in 10 minutes.

If the current is carried from higher to lower potential, equations (18), (20), and (20 *a*) give the electrical energy expended: there is a decrease of electrical energy and a development of other forms of energy. If the current is carried from lower to higher potential, there is an increase in the electrical energy, and these equations give the external energy which must be supplied.

The student should now understand why the electrical energy obtained by a charged rod, an electrophorus, or an electrostatic machine is relatively small. While one is able to develop high potentials with one of these, the quantity of electricity that can be produced is so small that the product of the voltage and the quantity is relatively small.

419. Power expended in a part of a circuit. Power has been defined as the rate of doing work. When the work is done at a uniform rate,

$$\text{Power} = \frac{\text{work}}{\text{time}}.$$

If both sides of equation (20) of the last section are divided by the time, we have the equation

$$\frac{\text{Work}}{\text{Time}} = \text{current} \times \text{voltage},$$

or

$$\text{Power} = \text{current} \times \text{voltage}. \tag{21}$$

In practical units (see equation (20 *a*))

$$\text{Power} = \frac{\text{joules}}{\text{seconds}} = \text{amperes} \times \text{volts}. \tag{22}$$

Since 1 *watt*, the unit of power, is equal to 1 joule per second,

$$\text{Watts} = \text{amperes} \times \text{volts}. \tag{23}$$

For example, when the difference of potential of the terminals of an incandescent lamp is 100 volts and the current 0.5 ampere, the power supplied is 50 watts.

It is customary to rate the sizes of incandescent lamps by the *power* used when they are connected to wires which have the difference of potential marked on the lamp. For example, a 50-watt, 110-volt lamp is one that will use 50 watts of power when connected to a 110-volt circuit. In a similar way it is customary to rate motors and other machines by the number of watts of power used under specified conditions.

For large amounts of power it is customary to use *kilowatts* instead of *watts*.

$$1 \text{ kilowatt} = 1000 \text{ watts}.$$

$$746 \text{ watts} = 1 \text{ H.P.}$$

The advanced student will learn that it is the practice in *alternating-current* measurements to use an average value of *amperes* and *volts* instead of the true, or instantaneous, values. When such averages are used, the formulas given for work and power do not always hold; an additional factor, called the *power factor*, must often be employed. However, these equations may be applied, without any appreciable error, to parts of alternating circuits that contain only resistance, such as incandescent lamps and heating-coils.

420. Other units for measuring work. The measurement of power in watts and kilowatts is so easily done in actual practice that it has become customary to determine the work by multiplying the power by the time. For example, when *watts* are multiplied by *seconds*, the work is given in *joules*. Because the second is so small a unit of time, the use of the *hour* as the unit is customary, and the work is computed by multiplying *watts* by *hours*. When this is done, the work is measured in *watt-hours*. When *kilowatts* are multiplied by *hours*, the work is measured in *kilowatt-hours*. The kilowatt-hour is used by electricians more than any other unit of work or energy. In nearly all towns and cities electrical energy is sold at so much a kilowatt-hour. The price varies, depending on the size of the electrical plant, on its location, and also on the amount of electricity the consumer uses.

421. Practical examples. 1. How much does it cost to run an electric iron for 3 hr. if the meter rate is 10¢ per kilowatt-hour and if the iron takes 5 amperes at 110 volts?

$$\text{Power} = 5 \times 110 = 550 \text{ watts.}$$

$$\text{Work} = 550 \times 3 = 1650 \text{ watt-hours}$$

$$= 1.65 \text{ kilowatt-hours.}$$

At 10¢ per kilowatt-hour this will cost 16½¢.

2. What does it cost per working-day of 8 hr. to operate a motor which supplies 25 H.P., assuming that the efficiency of the motor is 90 per cent and that the cost of the electrical energy is 6¢ per kilowatt-hour?

$$25 \text{ H.P.} = 25 \times 746 \text{ watts} = 18{,}650 \text{ watts} = 18.65 \text{ kilowatts.}$$

Since the motor has an efficiency of 90 per cent, it must receive power at the rate of

$$18.65 \div 0.90 = 20.72 \text{ kilowatts.}$$

$$\text{Work done in 1 day} = 20.72 \times 8 \text{ kilowatt-hours} = 165.76 \text{ kilowatt-hours.}$$

Hence $\text{Cost} = 165.76 \times \$0.06 = \$9.95.$

422. Work done in a complete circuit: electromotive force. The student has now learned that in computing the energy expended in a lamp, an electric stove, or anything through which an electric current is flowing, the quantity of electricity is multiplied by the difference of potential of the two terminals of the thing in which the work is done.

Obviously this method of computing work applies only to a *part* of a circuit, not to a *complete* one (a little thought will show one that the term *difference of potential* has no definite meaning when applied to a complete circuit).

In computing the electrical energy which has been converted into other forms of energy in a *complete* circuit, the term *electromotive force*, or *E.M.F.*, is used.

The electromotive force of a circuit is numerically equal to the work done when a unit quantity of electricity is carried around the complete circuit.*

If the work done in carrying 1 unit of electricity is equal to the electromotive force, then the work done in carrying Q units is

$$\text{Work} = Q \times \text{E.M.F.} \tag{24}$$

The electromotive force is measured in *volts*; hence the work in a complete circuit may be indicated by the equation

$$\text{Joules} = \text{coulombs} \times \text{volts} \tag{25}$$

$$= \text{amperes} \times \text{seconds} \times \text{volts}. \tag{26}$$

If both sides of equation (24) are divided by the time t,

$$\frac{\text{Work}}{t} = \frac{Q}{t} \times \text{E.M.F.},$$

or

$$\text{Power} = I \times \text{E.M.F.} \tag{27}$$

Equation (23), in practical units, applies:

$$\text{Watts} = \text{amperes} \times \text{volts}.$$

In any circuit there must always be some source of energy, usually located in a cell or in a dynamo; and since the electromotive force is used in computing the total energy supplied, it is customary to associate the electromotive force with this source of energy. Hence the electromotive force is said to be in the cell or in the dynamo. For example, the electromotive force of a dry cell is given as 1.5 volts. What is meant by this statement is that the work this cell can do in sending a unit charge (1 coulomb) around the entire circuit is equal to 1.5 joules. If the cell produces a current, I, which flows for t seconds, the quantity, or charge, is It, and the work done by the cell is equal to the electromotive force times the quantity, or $1.5\ It$.

---

* The term *work* as used here is equal to the amount of electrical energy which is converted into other forms of energy. It may be regarded, however, as referring either to the *output* of the circuit or to the *intake*, but not to both.

Electromotive force has a misleading name, for it is not a force but a work factor; when multiplied by *quantity*, it gives *work*. If it were a force, it would have to be multiplied by a *distance* to give work. The name had its origin in the analogy of an electric current to the flow of a liquid. To keep water in motion a mechanical force is necessary, and in the case of electricity the electromotive force was taken to be the moving force. While the name is not accurate, the beginner often finds the point of view helpful.

423. Summary. *Current* of electricity is the rate at which *quantity* of electricity passes a given place.

If the current is uniform, the quantity Q passing in a time t is, by equation (10),

$$Q = It;$$

if the flow is not uniform, this equation is true if I is the average current.

A summary of the facts regarding potential and potential differences is given in section 415. The *difference of potential* between any two points is equal to the work done in transferring *unit* quantity of electricity from the point of lower potential to that of higher potential. The work, W, done when a quantity, Q, is transferred from one place to another, through a difference of potential of $V_2 - V_1$, is, by equation (18), $W = Q\ (V_2 - V_1)$.

In practical units

$$\text{Joules} = \text{coulombs} \times \text{volts}.$$

When a current flows through conductors of any shape and made of any kind of material, there is always a loss, or fall, of potential. The resistance of the conductor is so defined that, by equation (12),

$$\text{Fall of potential} = \text{current} \times \text{resistance}.$$

In practical units, by equation (13),

$$\text{Loss in volts} = \text{amperes} \times \text{ohms}.$$

This equation is one form of Ohm's law.

Experiments show that the resistances of conductors made of the common metals are independent of the size of the current.

The *electromotive force* of a circuit is equal to the energy supplied to the circuit when unit quantity of electricity flows around the *complete* circuit. When Q units flow, the work done is, by equation (24),

$$W = Q \times \text{E.M.F.}$$

In a part of a circuit, by equation (21),

$$\text{Power} = I \times V,$$

and, by equation (20), $\text{Work} = I \times V \times t,$

where V is the difference of potential between the terminals of that part of the circuit.

In a complete circuit, by equation (27),

$$\text{Power} = I \times \text{E.M.F.}$$

and $\text{Work} = I \times \text{E.M.F.} \times t.$

In general, by equation (23),

$$\text{Watts} = \text{amperes} \times \text{volts.}$$

$$\text{Joules} = \text{watt-seconds} = \text{amperes} \times \text{volts} \times \text{seconds} \tag{28}$$

$$\text{Watt-hours} = \text{watts} \times \text{hours} \tag{29}$$

$$= \text{amperes} \times \text{volts} \times \text{hours} \tag{30}$$

$$\text{Kilowatt-hours} = \text{kilowatts} \times \text{hours} \tag{31}$$

$$= \frac{\text{amperes} \times \text{volts}}{1000} \times \text{hours} \tag{32}$$

PROBLEMS

1. The resistance of an electric toaster is 20 ohms. What will be the current when it is connected to two wires having a difference of potential of 106 volts?

2. An incandescent lamp has a current of 0.5 ampere when the voltage across it is 100 volts. What is its resistance?

3. A copper trolley wire has a resistance of 0.5 ohm per mile. A street car 2 mi. from the power house is using 80 amperes. What voltage is lost on the line?

4. 60 street lamps are connected in series. Each lamp requires 6 amperes and has a resistance of 3 ohms. (*a*) How much voltage is required for each lamp? (*b*) How much for all the lamps?

5. A current of 6 amperes flows for 4 min. through a resistance of 12 ohms. Compute (*a*) the applied voltage and (*b*) the total energy supplied in joules.

6. A house is connected to a power circuit in the alley by a line having a resistance of 0.5 ohms. If the voltage in the alley is 108 volts, what will it be in the house when a current of 10 amperes is flowing into the house?

7. A house is connected to a power circuit by 400 ft. of wire which has a resistance of 1.6 ohms per 1000 ft. What is the voltage drop in the wires (*a*) if the current is 5 amperes? (*b*) if the current is 10 amperes?

8. A headlight lamp on an automobile has 3 amperes at 6 volts. How many watts does it use?

9. (*a*) Find the resistance and the normal current for a lamp labeled "110 volts, 25 watts"; (*b*) for one labeled "60 volts, 25 watts"; (*c*) for one labeled "110 volts, 100 watts."

10. An electric toaster has a resistance of 20 ohms. How much power does it use when connected to a 100-volt circuit?

11. An incandescent lamp at 100 volts takes 0.5 ampere. Find the cost per hour of running this if the rate is 12 ¢ per kilowatt-hour.

12. Find the cost per hour of running an electric stove which is labeled "8 amperes, 110 volts" if the charge is 10 ¢ per kilowatt-hour.

13. An electric fan running on a 100-volt circuit uses 40 watts of power. At 10¢ per kilowatt-hour, how much does it cost to run the fan 8 hr.?

14. A dynamo, E.M.F. of 125 volts, is connected to a distant motor by wires having a resistance of 0.7 ohm. If the current is 40 amperes, (*a*) how much is the voltage drop along the wires? (*b*) what is the total power developed? (*c*) what is the power wasted on the wires?

CHAPTER XXIX

HEATING EFFECTS; JOULE'S LAW

The electric heater, 424. Electric fuses, 425. The incandescent lamp, 426. Joule's law, 427. Determination of the mechanical equivalent of heat by the use of Joule's law, 428. A theoretical proof of Joule's law, 429. A special case of electric heating; constant voltage, 430. Losses along transmission lines, 431. Why electrical energy is transmitted at high potentials, 432. Summary, 433.

424. The electric heater. The heating effects of an electric current are well known. In the ordinary incandescent lamp the current raises the temperature of the filament to "white" heat. Electric irons, toasters, and other heating devices are now common household appliances. While it is desirable in many cases for a current to develop heat, yet there are instances where this heating effect is a source of trouble and great danger. For example, the overheating of electric motors sometimes results in the destruction of the insulation of the wires, thus rendering the whole machine worthless. The principles that will be explained in this chapter apply to all cases: to those where heating is desirable, as well as to those where it is to be eliminated as far as possible.

Coils of wire are used as conductors in some types of electric heaters, and in others long, flat, metallic ribbons. As these have to be heated by the current to a bright red, they must have a rather high resistance; that is, they must be poor conductors, for good conductors are not so readily heated. In addition, the metal which is heated by the current must be one that is capable of being heated to a high temperature without oxidizing, or suffering some other form of deterioration. There are several kinds of alloys on the market that are excellent for heating-coils; for example, the nickel-chromium alloys are well suited for coils which are to be exposed to the air. Different alloys are usually known by trade names, such as *chromel*, *nichrome*, *excello*, *superior*, *climax*, and *advance*.

Two types of electric cookers are now in use: the electric range and the fireless cooker. The fireless cooker is similar to the ordinary type, except that it contains coils which are heated by an electric current. When the temperature reaches a certain value, a thermostat (sect. 194), which is mounted inside the cooker, will automatically open the circuit; when the temperature becomes too low, it will close the circuit. In general appearance the electric range resembles a gas or coal range. It shows some marked improvements in design, however, especially in the oven. This is usually surrounded by some thermally insulating material to prevent escape of heat, and a thermostat is frequently attached. It is also common to have a clock control, so that the electric circuit will be cut off at some predetermined time. Electrical heating usually costs more than other methods, but the advantages in the convenience of the arrangement are quite obvious.

425. Electric fuses. One of the common safety devices employed on electric circuits is the electric *fuse*. A short piece of aluminum wire, or one made from an alloy of lead and tin, which melts at a low temperature, is inclosed in a cartridge, or plug, and so connected to the electric circuit that all the current will flow through this piece of wire. When for any reason the current exceeds a certain value, the heat developed in this wire will melt it and thus break the circuit. Fuses are commonly employed in houses to prevent the current in the wires from reaching so large a value as to overheat the wires and thus lead to danger of fire. They are also used to protect instruments, small motors, and other pieces of apparatus which might be injured by a large current.

426. The incandescent lamp. The incandescent lamp is too well known to need any detailed explanation or description. Everyone knows that the electric current heats the filament to such a high temperature that it gives out light. However, only a few appreciate the remarkable work that has been done to bring this lamp to its present high development. One of the difficulties connected with getting a highly efficient source of light is that light usually constitutes a very small percentage of the total energy radiated from a hot body. It is known that the higher the temperature of the radiating body, the larger is the percentage of the total radiant energy which is in the form of light. Hence the aim has been to make a filament which will stand a higher temperature. The refractory metals tantalum and tungsten have replaced the carbon filaments because they can be heated to a higher temperature without volatilizing. Tungsten has the highest melting point of all the metals. It is ordinarily very hard; but methods of treating it have been developed which make it quite ductile, so that now fine wires are easily drawn. Nearly all lamp filaments are now made of tungsten.

But even tungsten may be volatilized by an electric current. When an inert gas—usually argon, or some mixture of which argon is the chief constituent—is put inside the lamp, a much higher temperature can be maintained without such a rapid disintegration of the filament as would occur in a vacuum. These gas-filled lamps, which were invented by Langmuir, owe their relatively high efficiency to the fact that the filament can be heated to a very high temperature. They are made in many sizes—from small automobile lamps up to those using ten kilowatts of power; even larger sizes have been made in the research laboratories.

427. Joule's law. In a great many cases it is important to be able to compute the amount of energy converted into heat when a current flows through any wire or any electrical apparatus. The quantitative law was first found experimentally by Joule. He discovered that the heat energy developed was proportional (1) to the *resistance* of the conductor, (2) to the *square* of the *current*, and (3) to the *time*.

The heat energy developed by a current I, measured in amperes, in a conductor of resistance R ohms, may be expressed by the equation

$$\text{Heat energy in joules} = I^2Rt. \qquad (33)$$

From the known relation between joules and calories,

$$\textbf{Energy in calories} = \frac{1}{4.187} I^2Rt, \qquad (34)$$

or (approximately)

$$W = 0.24\, I^2Rt \text{ cal.} \qquad (35)$$

The *rate* of development of heat, or the *power* expended in heat, is found by dividing equation (33) by the time t:

$$P = I^2R \text{ watts,} \tag{36}$$

$$P = \frac{I^2R}{4.187} \text{cal./sec.,} \tag{37}$$

or (approximately) $$P = 0.24\, I^2R \text{ cal./sec.} \tag{38}$$

Equations (33) and (35) are different statements of Joule's law.

As an example of the use of Joule's law consider the following case: Coils of wire having a resistance of 2 ohms are placed in series with the arc lamp of a projection lantern, to prevent the current from becoming excessively large. How much power is *lost* when the current is 15 amperes? According to Joule's law the power lost is I^2R watts, or $15^2 \times 2 = 450$ watts. This is equal to $0.24 \times 450 = 108$ calories per second.

Joule's law is entirely general, applying to all cases. It enables one to compute the amount of heat developed by a current in a motor or in any other type of electrical apparatus which has resistance. In most electrical apparatus the development of heat is to be avoided; but if a certain current must flow, the heat loss can be decreased by making the resistance small. In other cases the heat loss can be reduced by using smaller currents.

428. Determination of the mechanical equivalent of heat by the use of Joule's law. An electrical method is not only the most convenient method for determining the numerical value of the mechanical equivalent of heat but also the most accurate one. A mechanical method for obtaining this important constant was described in Chapter XIX. As stated in that chapter, the relation between *joules* and *calories* is

$$W = JH, \tag{39}$$

where W is the numerical value of a certain quantity of energy measured in joules, H the numerical value of the same quantity measured in calories, and J the mechanical equivalent of heat.

If a coil of wire is immersed in a calorimeter containing water, the energy supplied to it in t seconds is, by Joule's law, equal to I^2Rt joules, where I is the current in amperes flowing through the coil, and R is its resistance in ohms.

If the mass of the water and the water equivalent of the calorimeter and coil of wire are determined, and if the change in temperature is observed, the number of calories supplied, H, can be computed. The numerical value of the same amount of energy in both joules and calories can thus be obtained. This gives sufficient data to determine the value of J in equation (39).

429. A theoretical proof of Joule's law. Although the law which bears Joule's name was determined experimentally, it is possible, by using the principle of the conservation of energy, to prove it theoretically. The proof is simple.

When a current, I, flows through a resistance—say of value R—there is a difference of potential, V, between the two ends of this resistance. Equation (20 *a*), sect. 418, gives, as the energy delivered to this part of the circuit in a time t,

$$W = IVt \text{ joules.}$$

But, by Ohm's law, $V = IR.$

Substituting this value for the energy supplied, we have

$$W = I(IR)t = I^2Rt \text{ joules.}$$

None of this energy, so far as we know, goes into any other form of energy than that of *heat*. Hence we may assume, in accordance with our principle of conservation of energy, that the amount of electrical energy supplied is equal to the heat energy developed, or

$$\text{Heat energy} = I^2Rt \text{ joules.}$$

This proof is not entirely general. The form of Ohm's law which is used does not, as will be explained later, apply to all cases; but it is only a matter of detail to use a more general law, and thus to derive Joule's law for all cases.

430. A special case of electric heating; constant voltage. The *rate* of heating has been shown to be proportional to I^2r. One must not infer from this that the rate is *always* proportional to the resistance. There are cases where the rate will vary inversely as the resistance; that is, the larger the rate the smaller the resistance. For example, take the case of the ordinary lighting circuit, where the voltage between the two wires is, say, 100 volts. If a lamp of 100 ohms is connected to these wires, the current through the lamp will be 1 ampere, and the rate of heating in watts will be $1^2 \times 100$

= 100 watts. Now suppose a lamp of 200 ohms is connected to the wires. The current in the lamp will be 0.5 ampere, and the rate of heating will be $0.5^2 \times 200 = 50$ watts. The lamp of the greater resistance will receive heat at a lower rate. The student should have little difficulty in seeing that this is due to the fact that when the resistance is larger, as in this case, the current is smaller. The fact that the heating effect varies inversely as the resistance is true *in all branched, or divided, circuits.* The branch having the greater resistance has the smaller current and a smaller rate of heating.

431. Losses along transmission lines. All wires used in the transmission of electrical energy have resistance. For example, the resistance of one mile of copper wire, No. 0 American wire gauge, is about 0.5 ohm. This is a rather large-sized wire having a diameter a little greater than that of the average lead pencil. Such wire is expensive, not only to buy but also to put on the poles; for in 100 feet of it the copper alone weighs over 30 pounds. The power lost in a line of this wire, used to convey electrical energy from a power house to a mill 2 miles away, will be computed. The complete circuit would require 4 miles of wire having a total resistance of about 2 ohms. If the current flowing in this wire were 20 amperes, the power wasted in heat along the line would be $I^2R = 20^2 \times 2 = 800$ watts, or over 1 horse power.

432. Why electrical energy is transmitted at high potentials. Consider the case of a 2-wire transmission line, 3 miles long, of No. 0 copper wire. Six miles of wire would be needed. At 0.5 ohm per mile, it would have a total resistance of 3 ohms. According to Joule's law—the I^2R law—the power in watts that is wasted in heat for three different currents is as follows:

For 10 amperes 300 watts.
For 50 amperes 7500 watts.
For 100 amperes 30,000 watts.

In the case of the 10-ampere current the loss might be regarded as permissible, but not so in the other cases.

One must keep in mind that it is *energy* which is to be transmitted. The power, or energy transmitted per second, is equal to the current times the voltage between the wires, or

$$\text{Power} = \text{current} \times \text{voltage}.$$

A glance at this equation should show that the power may be large and the current small *if the voltage is made large.* Hence large losses can be avoided by keeping the current small and by using high voltages when large amounts of power are desired. A few cases are worked out to make this entirely clear.

Suppose that the power delivered to a line of 3 ohms in each of the following cases is 50 kilowatts:*

1. The difference of potential of the wires at the power-house end of the line is 500 volts.

$$\text{Power} = 50 \text{ kilowatts} = 50{,}000 \text{ watts.}$$

$$\text{Voltage} = 500.$$

$$\text{Current} = 100 \text{ amperes.}$$

$$I^2R \text{ (loss in line)} = 100^2 \times 3 = 30{,}000 \text{ watts,}$$

a loss of 60 per cent.

2. The difference of potential at the power house is 10,000 volts.

$$\text{Power} = 50 \text{ kilowatts} = 50{,}000 \text{ watts.}$$

$$\text{Voltage} = 10{,}000.$$

$$\text{Current} = 5 \text{ amperes.}$$

$$I^2R \text{ (loss)} = 5^2 \times 3 = 75 \text{ watts, a small loss.}$$

For distribution of power for light-circuits in cities and towns, about 2200 volts are commonly used; but where the distances are much greater, higher voltages are employed. For example, the electrical power from the hydroelectric plant on the Mississippi River at Keokuk is delivered to the transmission lines at about 110,000 volts.

---

* To avoid complications these examples have been treated as though they were cases of the transmission of energy by direct currents. If alternating currents were used, and if average values were given for the quantities concerned, the power equation would be (see last paragraph of section 419)

$$\text{Average power} = \text{average voltage} \times \text{average current} \times \text{power factor.}$$

The power factor is different for different circuits, varying from about 0.8 to 0.95; in some cases it may be as low as 0.6. When the circuit is used almost entirely for lighting purposes, the power factor may be regarded, in all rough computations, as equal to 1.0.

433. Summary. Joule's law states that the heat developed in a conductor by an electric current is, by equation (33),

$$\text{Heat energy in joules} = I^2Rt,$$

or, by equation (35),

$$\text{Heat energy in calories} = 0.24\ I^2Rt \text{ (approximately)}.$$

The power, P, expended in heat is, by equation (36),

$$P = I^2R \text{ watts},$$

or, by equation (38),

$$P = 0.24\ I^2R \text{ calories per second.}$$

When the terminals of any piece of apparatus are maintained at a constant voltage, the rate of heating varies inversely as the resistance. When two circuits are connected in parallel, the greater heat will be developed in the one with the smaller resistance.

The energy losses along transmission lines are due to the heat developed in the conductors. Equation (36) gives the power loss.

On account of the fact that the heating of conductors depends on the square of the current, and thus increases very rapidly as the current increases, it is customary to transmit power at high voltage. When the voltage is high, the current may be relatively small; for the power is always proportional to the product of current and voltage.

PROBLEMS

1. A current of 20 amperes flows over a transmission line which has a resistance of 0.5 ohm. What is the rate of loss of heat in watts?

2. The heating-coils of an electric iron have a total resistance of 18 ohms. When the iron is connected to supply mains which have a difference of potential of 108 volts, how many calories will be developed per minute?

3. A heating-coil to which 500 watts is supplied is placed in 1000 gm. of water. How long will it take to raise the temperature of the water from 20° C. to 100° C. if all loss of heat to the surroundings is negligible?

4. Calculate the number of calories produced when a 25-watt lamp is run for an hour.

5. An electric toaster has a resistance of 20 ohms. (*a*) How much power in watts does it use when connected to a 100-volt circuit? (*b*) How many calories will it develop in 30 min.?

6. Heat energy is supplied electrically at the rate of 75 cal./min. (*a*) Find the number of watts supplied. (*b*) If the E.M.F. is 2 volts, find the current.

7. A coil of wire having a resistance of 20 ohms is immersed in a vessel containing 300 gm. of water at 20° C. The wire carries a current of 2 amperes for 10 min. If all the heat generated is used to warm the water, what is the final temperature of the water?

8. A transmission line has a resistance of 2 ohms. (*a*) How much power is lost when a current of 10 amperes flows in the line? (*b*) a current of 100 amperes?

9. How much will it cost to raise the temperature of 10 gal. of water (1 gal. has a mass of 3800 gm.) from 10° C. to 60° C. by electrical energy which is supplied at 10¢ per kilowatt-hour?

10. 50 kilowatts of electrical power is delivered to a long transmission line which has a resistance of 5 ohms. (*a*) What will be the power wasted in the line if the power is delivered at 10,000 volts? (*b*) at 1000 volts?

CHAPTER XXX

RESISTANCE; GENERAL FORM OF OHM'S LAW

Summary of the principal facts about the resistance of conductors, 434. Rheostats; resistance boxes, 435. Measurement of resistance; the Wheatstone-bridge method, 436. Resistances in series, 437. Resistances in parallel, 438. Proof of the law of parallel resistances, 439. Law of divided circuits; shunts, 440. Ohm's law; loss of potential due to resistance, 441. Ohm's law applied to a complete circuit; electromotive force, 442. Ohm's law applied to groupings of cells, 443. General form of Ohm's law, 444. Limitations of Ohm's law, 445. Counter electromotive force in circuits, 446. Summary, 447.

434. Summary of the principal facts about the resistance of conductors. 1. The resistance of a conductor has the same numerical value for large currents as for small currents, provided the temperature is not changed. It is this fact that makes Ohm's law so useful. If the resistance of a given piece of electrical apparatus has once been found, its value is known for other occasions.

(2) The resistance of a conductor varies inversely with the cross-sectional area.

(3) The resistance of a conductor is directly proportional to its length.

(4) The resistance of a conductor depends on the material (see section 413 for the relative resistances of a number of metals). Impurities in a metal usually tend to increase its resistance. Since alloys of metals have much greater resistance than pure metals, wires made of alloys are commonly used where a considerable resistance is needed; for example, in the coils of electric heaters. Metals have much lower resistance than liquids.

(5) The resistance of metallic conductors usually increases as the temperature rises. The resistance of liquids, such as those used in cells, decreases as the temperature rises. For example, the resistance of a storage battery is higher at low temperatures. The re-

sistance of most of the pure metals increases at about the same rate—about 4 per cent for an increase of 10° C. Alloys usually have a small rate of change. In the case of a few alloys the resistance remains practically constant over a wide range of temperature. This is one of the reasons that certain alloys are used as conductors in many types of electrical instruments. If α, the "temperature coefficient," and R_0, the resistance at 0° C., are known, the resistance R_t at a temperature t can be computed from

$$R_t = R_0(1 + \alpha t).$$

TEMPERATURE COEFFICIENTS OF RESISTANCE OF METALS PER DEGREE CENTIGRADE

Aluminum	0.0042	Manganin	0.00001
Copper (hard-drawn)	0.00408	Platinum	0.0038
Iron (hard steel)	0.0016	Silver	0.0040
Iron (soft steel)	0.0042	Tungsten	0.0051

435. Rheostats; resistance boxes. In the laboratory it is often necessary to insert in a circuit resistance which can be increased or decreased in order to obtain a current of the desired magnitude.

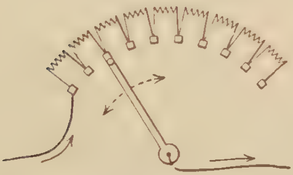


FIG. 259

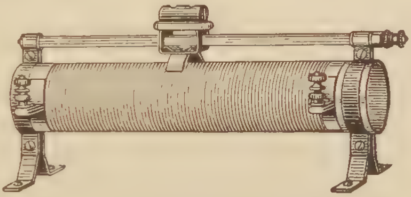


FIG. 260

For this purpose there are many types of commercial *rheostats*. They usually contain coils of wire with switches or sliding contacts so that different portions can be used (see Figs. 259 and 260).

Resistance boxes differ from rheostats in that the length and size of the wire in the coils are so chosen that the resistances have definite values, while in a rheostat the values of the resistances are usually not given. The coils in resistance boxes are usually made of fine wire which will not carry much current, while rheostats are much more rugged in construction.

436. Measurement of resistance; the Wheatstone-bridge method. The accurate measurement of the resistance of a coil of wire, or of some instrument, plays an important part in electrical measurements. Hence the theory of one common method of determining resistance will be given. This is usually referred to as the *Wheatstone-bridge* method.

If four resistances, represented by r_1, r_2, r_3, and x, are connected to a galvanometer G and the cells B, as shown in Fig. 261, and if these resistances are adjusted until no current flows through the galvanometer, then, as will be proved later, the following relation is true:

$$\frac{x}{r_1} = \frac{r_3}{r_2}.$$

Let x be the unknown resistance to be measured. This equation may be written

$$x = \frac{r_1}{r_2} r_3. \qquad (40)$$

When the resistance r_3 and the ratio r_1/r_2 are known, x is readily computed. In practice, r_1 and r_2 have some convenient value, and the resistance r_3 is varied until the "bridge" is balanced; that is, until no current flows through the galvanometer. Or r_3 may be kept fixed, and the ratio r_1/r_2 changed.

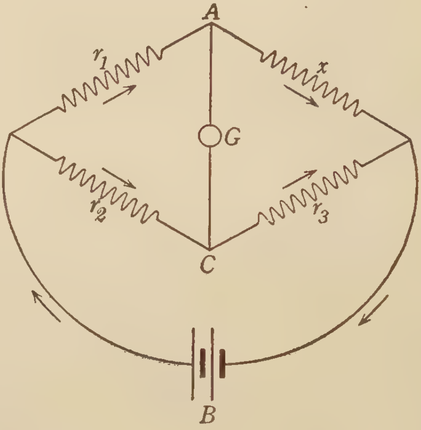


FIG. 261

It is not difficult to derive the preceding law. The fall in potential when a current i_1 flows through the resistance r_1 is i_1r_1, and the fall when a current i_2 flows through r_2 is i_2r_2. If there is no current through the galvanometer, the two points marked A and C in the figure must be at the same potential, or the fall of potential through r_1 must be the same as that through r_2. Hence

$$i_1r_1 = i_2r_2. \qquad (41)$$

In a similar way, the fall of potential through the resistance x must be the same as that through r_3; hence

$$i_x x = i_3 r_3.$$

But the current i_x must be equal to i_1, since all the current flowing through r_1 goes through x. Similarly, the current i_3 must be equal to i_2. Substituting these in the last equation, we have

$$i_1 x = i_2 r_3. \tag{42}$$

If equation (42) is divided by equation (41),

$$\frac{i_1 x}{i_1 r_1} = \frac{i_2 r_3}{i_2 r_2},$$

or

$$\frac{x}{r_1} = \frac{r_3}{r_2}.$$

Hence (equation (40)) $x = \frac{r_1}{r_2} r_3.$

Various types of bridges based on this relation are in use. The student will probably have an opportunity of using one or more types in the laboratory.

437. Resistances in series. Fig. 262 shows, in a conventional way, a cell connected in series with three resistances. The resistances of these are marked as 2, 4, and 6 ohms respectively. The total resistance connected to the cell is the sum of these, or 12 ohms.

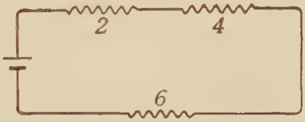


Fig. 262

When resistances are connected in series, the combined resistance is the arithmetical sum of the separate resistances.

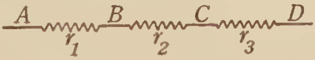


Fig. 263

This theorem may be proved in the following manner: When a current I flows through the resistances shown in Fig. 263, the fall of potential through the resistance r_1 is Ir_1, through r_2 is Ir_2, and through r_3 is Ir_3. The total fall of potential between A and D is

$$Ir_1 + Ir_2 + Ir_3.$$

But, by the definition of resistance, this must be equal to IR, where R is the total resistance between A and D. Hence

$$IR = Ir_1 + Ir_2 + Ir_3,$$

or

$$R = r_1 + r_2 + r_3, \tag{43}$$

which is the relation to be proved.

438. Resistances in parallel. In Fig. 264 are shown diagrammatically three resistances in *parallel*. In this case the combined resistance must be *less* than the resistance of any one of them. To illustrate this let us first consider the case of the flow of water through pipes. Suppose the two large mains A and B (Fig. 265) are connected only by the pipe C. When the pipe D is added, another path is offered for the flow of water; hence more water will flow, and the *resistance* to the flow is decreased. In a similar way, if r_1 of Fig. 264 is the only resistance, adding the resistances r_2 and r_3 offers additional paths for the current and lowers the resistance.

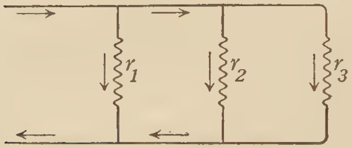


FIG. 264

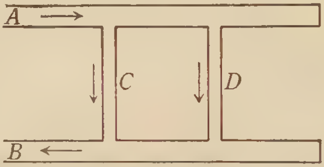


FIG. 265

The law for computing the total resistance, R, of three resistances r_1, r_2, and r_3 connected in parallel is

$$\frac{1}{R} = \frac{1}{r_1} + \frac{1}{r_2} + \frac{1}{r_3}. \tag{44}$$

For example, let $r_1 = 10$ ohms, $r_2 = 5$ ohms, and $r_3 = 20$ ohms.

Then

$$\frac{1}{R} = \frac{1}{10} + \frac{1}{5} + \frac{1}{20} = \frac{2+4+1}{20};$$

$$\frac{1}{R} = \frac{7}{20},$$

or

$$R = \frac{20}{7} = 2.86 \text{ ohms.}$$

If the resistances r_1, r_2, and r_3 are all equal, it is easily proved from equation (44) that

$$R=\frac{r}{3},$$

where r is the resistance of one of the three equal resistances.

439. Proof of the law of parallel resistances. The law of parallel resistances (see equation (44)) will now be proved. Let the current be flowing from A to B (Fig. 266). Then A is at a higher potential than B by an amount of, say, V volts. This amount must be the same for each of the three branches. By Ohm's law the current through the branch the resistance of which is r_1 is given by the equation

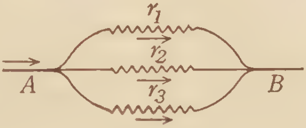


FIG. 266

$$i_1=\frac{V}{r_1}. \tag{45}$$

In a similar way the current through the second branch is

$$i_2=\frac{V}{r_2}, \tag{46}$$

and through the third,
$$i_3=\frac{V}{r_3}. \tag{47}$$

The total current, I, through the system is

$$I=i_1+i_2+i_3. \tag{48}$$

Substituting in equation (48) the values of i_1, i_2, and i_3, we obtain

$$I=\frac{V}{r_1}+\frac{V}{r_2}+\frac{V}{r_3},$$

or

$$I=V\left(\frac{1}{r_1}+\frac{1}{r_2}+\frac{1}{r_3}\right). \tag{49}$$

But if R is the resistance of the combination, we have, from the definition of resistance,

$$I=V\frac{1}{R}. \tag{50}$$

Comparing equations (49) and (50), we see that (equation (44))

$$\frac{1}{R}=\frac{1}{r_1}+\frac{1}{r_2}+\frac{1}{r_3}.$$

440. Law of divided circuits; shunts. When a current (Fig. 267) divides,—part flowing in one branch and the rest in another,—the branch containing the smaller resistance will carry the larger part of the current. The law is that the ratio of the currents in the two branches is inversely proportional to the ratio of the resistances of the branches; or, expressed in an equation,

$$\frac{i_1}{i_2}=\frac{r_2}{r_1}, \qquad (51)$$

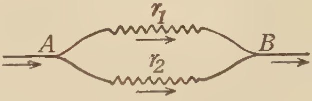


FIG. 267

where i_1 and r_1 represent the current and the resistance of one branch, and i_2 and r_2 the current and the resistance of the other branch. This law is readily proved. Let V represent the potential difference between the points A and B where the two branches join (Fig. 267). Since V is equal to the voltage drop through each branch,

$$V=i_1r_1 \qquad (52)$$

and

$$V=i_2r_2. \qquad (53)$$

Equating the right-hand sides of equations (52) and (53), we have

$$i_1r_1=i_2r_2.$$

Dividing both sides by i_2r_1, we have (equation (51))

$$\frac{i_1}{i_2}=\frac{r_2}{r_1}.$$

It often happens that one wants only a part of a current to flow through an instrument. For example, if one has a galvanometer or an ammeter that is not capable of carrying a large current, a conducting wire or a coil of wire may be connected across the terminals of the instrument so that the current will divide, part going through the instrument and part through the wire. A conductor used in this way is called a *shunt*.

If i_1 is the current through the instrument, I the total current, r_1 the resistance of the instrument, and r_2 the resistance of the shunt, it will be proved that the current through the instrument is

$$i_1 = I \frac{r_2}{r_1 + r_2}.$$

By a well-known law of proportion, equation (51) becomes

$$\frac{i_1}{i_1 + i_2} = \frac{r_2}{r_1 + r_2}.$$

But

$$i_1 + i_2 = I,$$

or

$$\frac{i_1}{I} = \frac{r_2}{r_1 + r_2}.$$

Hence

$$i_1 = I \frac{r_2}{r_1 + r_2}. \qquad (54)$$

By the aid of this equation one can compute the current in a branch if the total current and the resistances of the two branches are known. For example, if the resistance of an ammeter, without its shunt, is 8.45 ohms, and the shunt alone has 0.05 ohm, the fraction of the total current that goes through the instrument is

$$\frac{0.05}{8.45 + 0.05} = \frac{1}{170}.$$

441. Ohm's law; loss of potential due to resistance. The form of Ohm's law that has been used again and again in this book must be regarded as only a special case of a more general form; for, as will be seen, it applies only to certain special cases. In the form

$$V = Ir \qquad (55)$$

V is the fall of potential when a current I flows through a resistance r. But there may be other causes for changes in the potential. The loss or fall of potential in a part of a circuit is (sect. 416) always equal to the *work* done when a unit quantity of electricity has passed through that part of the circuit. When there is only resistance in the circuit, this work all goes into heat energy in the conductors. In such cases equation (55) is strictly true. But there are other causes which may affect the work required in order to send a unit quantity through the circuit. Thus, suppose that the current is driving a motor. In this case, energy which does not go into heating the conductors is expended. Or if there is a storage battery being charged, energy is expended in the chemical reactions known as "charging." More work must be done in such cases in forcing the current through than when the circuit contains only resistance. Hence the total fall of potential, which is the work done for each unit charge, is *greater* than that given by the product Ir. On the other hand, there are places in every circuit where energy is being supplied to the current. For example, in going through an assisting cell or some other type of generator the current gains energy. In all such cases there are gains in the potential which are not taken care of by equation (55).

Equation (55) gives only the resistance loss of potential.

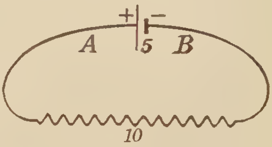


FIG. 268

An example will assist us in understanding this point. A cell, the internal resistance of which is 5 ohms, is connected to a resistance of 10 ohms (Fig. 268). We shall assume the current to be 0.1 ampere (the precise value of the current for this case is not important). What is the potential difference between the points A and B? If we apply equation (55) to the external part, the resistance 10, we obtain

$$V = 0.1 \times 10 = 1.0 \text{ volt};$$

if we apply it to the part of the circuit through the cell, we obtain

$$V = 0.1 \times 5 = 0.5 \text{ volt}.$$

It is obvious that the difference of potential between the points A and B cannot have two different values. Which one of these two is correct? If reference is made to the limitation of the law just pointed out, the student will see that the first one is correct. In the external part of the circuit we have only the resistance 10; but in the other part of the circuit, the part containing the cell, we have chemical reactions, which change the potential losses.

Hence the expression $V = Ir$ does not always give the total change in potential in a part of a circuit. It must be regarded as giving for any part of an electric circuit *that portion of the loss of voltage which is caused by resistance.*

442. Ohm's law applied to a complete circuit; electromotive force. When we are dealing not with a portion of a circuit but with a *complete* circuit, Ohm's law may be written

$$\text{Current} = \frac{\text{electromotive force}}{\text{total resistance}}, \tag{56}$$

where the resistance must be that of the entire circuit, including the cell or generator. As an example of this form of Ohm's law, let the electromotive force of the cell in Fig. 268 be 1.5 volts, its internal resistance 5 ohms, and the external resistance 10 ohms. The current is, then,

$$\frac{1.5}{5+10} = 0.1 \text{ ampere.}$$

This form of the law can be readily deduced from the one we have been using. To do this let us take a *closed* circuit and divide it into convenient parts, numbering the parts 1, 2, 3, and so on. For part 1 the loss in potential due to the resistance r_1 of that part may be written $V_1 = Ir_1$; for part 2, $V_2 = Ir_2$; for part 3, $V_3 = Ir_3$; and so on for all parts of the circuit. If we add up these equations, we have

$$\begin{aligned}\text{Total loss in potential} &= Ir_1 + Ir_2 + Ir_3 + \ldots \\ &= I \times \text{total resistance.}\end{aligned} \tag{57}$$

In a complete circuit there must be just as much gain in potential as there is loss. For if one could travel completely around such a circuit, measuring the potential as he goes, he would find when he came back to the starting-point that the potential was the same there as it was when he started. There are losses in the circuit, but there are also gains; for example, when the chemical action in a cell raises the potential. In a complete circuit

$$\text{Total loss} = \text{total gain of potential.}$$

If we call the *total gain* (or the *total loss*) the *electromotive force* (*E.M.F.*), and write for that quantity the letter E, then

$$\begin{aligned}E &= \text{total rise in potential} \\ &= \text{total loss in potential.}\end{aligned}$$

Hence, from equation (57),

$$E = IR, \tag{58}$$

where R is the total resistance. Obviously this equation may be written in the form of equation (56).*

Since the electromotive force is measured in *volts*, we have, as before,

$$\text{Volts} = \text{amperes} \times \text{ohms}.$$

However, the form of Ohm's law given in equation (55) differs from that given in equation (58), in that (55) applies to a *portion* of a circuit, while (58) applies to a *complete* circuit.

The electromotive force is here defined as equal to the *total* loss in potential. In section 422 the electromotive force of a cell was defined as equal to the work that the cell or other generator does in sending a unit charge around the entire circuit. When it is recalled that the loss of potential is equal to the work per unit charge, it is seen that there is complete agreement between the two statements.

443. Ohm's law applied to groupings of cells. *Cells in series.* When a number of cells or other sources of current are connected in *series*, the total electromotive force in the circuit is the algebraic sum of the electromotive forces of the individual cells. If the cells are so connected that they all assist, the resulting electromotive force is the arithmetical sum; but if any are connected so as to oppose the current, their electromotive force must be considered as negative in taking the sum.

If there are N cells in series in a battery, each cell having an electromotive force e and a resistance r, the current is given by

$$I = \frac{Ne}{R + Nr}, \tag{59}$$

where R is the resistance of the *external* part of the circuit.

Cells in parallel. When similar cells are connected in *parallel*, the resultant electromotive force is equal to that of one cell. The total resistance is obtained from the law of parallel resistances.

---

*The proof given here has for obvious reasons been limited to a simple type of circuit. It may, however, be extended without great difficulty to include a circuit containing motors or one used for charging storage batteries.

If N cells are connected in parallel, the electromotive force of each cell being e, and the internal resistance r, Ohm's law applied to the complete circuit is

$$I = \frac{e}{R + \frac{r}{N}}, \tag{60}$$

where R is, as before, the *external* resistance. This formula applies only to cells which are alike; that is, to those which have equal electromotive forces and equal resistances.

What is the advantage of connecting cells in parallel?

Cells in series-parallel. If the cells are connected in N_1 rows, each row consisting of N_2 cells in series, the rows being connected in parallel, Ohm's law becomes

$$I = \frac{N_2 e}{R + \frac{N_2 r}{N_1}}. \tag{61}$$

Why?

This law applies only when the cells are alike.

444. General form of Ohm's law. Equation (55),

$$V = Ir,$$

gives the fall, or loss, of potential which is caused by the resistance r in a *part* of a circuit, but does not include other losses in potential. On the other hand, equation (58),

$$\text{E.M.F.} = IR,$$

applies to a *complete* circuit where the *total* resistance is R.

Neither of these equations may be said to be the general law, as each applies to a particular case. The more general case can be derived as follows:

If between two points A and B on a circuit there is a cell or other generator, there is (1) a certain *fall* of potential due to the resistance, r, between A and B, and (2) a *rise* of potential due to the action of the generator.

The *fall* of potential due to resistance is equal to Ir, where r is the resistance of that part of the circuit. The *rise* is equal to the electromotive force. Therefore the resultant potential difference obtained (as measured by an electrometer or by a voltmeter) between the two points is

$$V = E - Ir. \tag{62}$$

Equation (62)* may be said to be the general form of Ohm's law.

Let us turn back to the problem of section 441. There we had a cell with a resistance of 5 ohms connected to an external resistance of 10 ohms, the current being 0.1 ampere.

The loss of potential in the cell was $0.1 \times 5 = 0.5$ volt; the loss in the external circuit was $0.1 \times 10 = 1.0$ volt. Therefore the total loss was 1.5 volts. In the cell there must have been a total rise of potential of 1.5 volts, or the electromotive force was 1.5 volts. As a check we may use equation (58), remembering that the total resistance is 15 ohms.

$$E = 0.1 \times 15 = 1.5 \text{ volts.}$$

In section 441 it was pointed out that the product Ir did not give the difference of potential of the terminals of the cell if r is the resistance of the cell. If we wish to compute the difference of potential at the terminals of the cell, we may use equation (62), or

$$V = 1.5 - (0.1 \times 5) = 1.0 \text{ volt.}$$

This agrees with the voltage as found by computing the fall of potential in the external circuit; that is,

$$0.1 \times 10 = 1.0 \text{ volt.}$$

The student should see that the 1.5 volts is the total *rise* in potential in the cell, and that the 0.1×5, or 0.5 volt, is the resistance loss in the cell, leaving a *net rise* in the cell of 1.0 volt.

Sometimes there is in part of a circuit an electromotive force which opposes the flow of the current. For example, in the charging of a storage battery the electromotive force of the battery opposes the flow of the current. In that case the fall of potential in the part of the circuit containing the storage battery is

$$V = E + Ir, \tag{63}$$

---

*From this equation both (55) and (58) are readily obtained as particular cases. If $E =$ zero (that is, if the part AB does not contain an electromotive force), we then have equation (55). Or if we imagine the points A and B to move away from the generator and along the circuit until they meet at a distant point, the difference of potential, V, will be zero, and r will be the total resistance. We then have equation (58).

where E is the *counter electromotive force* of the battery and r the resistance of that part of the circuit. For example, if the resistance of a storage battery is 0.2 ohm, the charging current 10 amperes, and the counter electromotive force of the battery 30 volts, a voltmeter connected to the terminals of the battery should read

$$30 + (10 \times 0.2) = 32 \text{ volts.}$$

445. Limitations of Ohm's law. In general, Ohm's law is not true for a current which is not constant in magnitude. When a circuit containing a battery is closed, a certain amount of time is required (usually very short) before the current reaches its final steady value. Ohm's law applies only after the current has reached its final value. Ohm's law does not always apply to an *alternating-current* circuit, such as is commonly used in lighting and power circuits. In a few special cases of alternating-current circuits this law may be used without appreciable error; for example, in the case of an alternating current through an incandescent lamp, or through the heating-coils used in electric heaters, such as irons and toasters. But in general one must be careful in applying Ohm's law to alternating-current machinery.

In the case of the conduction through gases and vapors, such as vacuum tubes and arc lights, Ohm's law is of little value. The resistance of gases and vapors is not constant, but changes with the current and the voltage.

446. Counter electromotive force in circuits. *When electrical energy other than that spent in developing heat is used in any part of a circuit, there is always present a counter electromotive force which opposes the flow of current.*

Energy must be supplied to an electric motor used for running machinery; that is, work must be done by the current to keep the motor running. This is always true, even for small motors, such as those used in running electric fans. Hence we must expect, from the statement made at the beginning of this section, to find present a counter electromotive force which will oppose the current. The presence of this force is easily shown by a simple experiment. If a small direct-current motor is connected to the power supply through a group of incandescent lamps, the brightness of the lamps will give an indication of the size of the current. It will be found that the current through the motor is much larger when there is no rotation in the motor than when it is allowed to run. The fact that the current is smaller when the motor is running shows that there is a counter electromotive force developed. In most motors the resistance is quite small; hence precautions must be taken, in starting a large motor, to avoid the excessively large current which otherwise would flow through it before it develops enough speed to give it a sufficiently large counter electromotive force. In the case of large direct-current motors, "starting" resistances are kept in the circuit until the motor has nearly reached full speed.

In charging a storage battery, work is done; hence there is a counter electromotive force in the cells which opposes the flow of the charging current.

In any circuit containing a counter electromotive force Ohm's law may be written in the form

$$I = \frac{E - e}{R}, \tag{64}$$

where E is the applied electromotive force, R the total resistance of the circuit, and e the counter electromotive force.

Proof of the principle. The statement that, when work other than that of developing heat is done, there is always a counter electromotive force is capable of a simple proof.

The total power expended in the circuit is, according to equation (27), equal to IE. The amount expended in heat is, according to equation (36), I^2R, where R is the total resistance. In addition to this, if work is being done, there is power expended where the work is done. This power is always proportional to the current; hence it is equal to some factor times the current, say Ie.

Equating the total power to the sum of the amounts expended in heat and in the other type of work,

$$IE = I^2R + Ie, \tag{65}$$

or Watts supplied = heat loss in watts + watts utilized.

This equation is a simple application of the principle of the conservation of energy.

Dividing equation (65) by I, and transposing,

$$IR = E - e,$$

or (equation (64)),

$$I = \frac{E - e}{R}.$$

Comparing this law with Ohm's law, we see that the factor e corresponds to a counter electromotive force.*

One important thing to be gathered from this is that the power delivered to a motor, or to a storage battery, or to any other apparatus where work of a mechanical or chemical character is being done, is always given by

$$\text{Effective power} = I \times \text{counter electromotive force.} \tag{66}$$

447. Summary.† When pieces of electrical apparatus are connected in series, the total resistance is the arithmetical sum of the resistances of the different parts.

The formula for the resistance of parallel connection is equation (44),

$$\frac{1}{R} = \frac{1}{r_1} + \frac{1}{r_2} + \frac{1}{r_3} + \cdots,$$

where R is the total resistance, and $r_1, r_2, r_3, \cdots$, the resistances of the separate branches.

In a circuit of two branches the current divides in inverse proportion to the resistance in the branches. Stated in equation form (equation (51)),

$$\frac{i_1}{i_2} = \frac{r_2}{r_1},$$

where i_1 and r_1 are the current and the resistance of one branch, and i_2 and r_2 the current and the resistance of the other branch. The current through one branch is given by equation (54),

$$i_1 = I\frac{r_2}{r_1 + r_2},$$

where I is the total current.

The expression Ir gives for a *part* of a circuit, of which the resistance is r, the fall, or loss, of potential that is caused by the resistance. Since there may be other causes which affect the change of potential, this expression gives the difference of potential when there is only resistance in that part of the circuit.

---

* Although the principle has been proved above for the case of direct currents, the advanced student will learn that it is true also for alternating currents. However, in that case conditions are more complicated, for other electromotive forces are usually introduced (see also section 487).

† A summary of the principal facts about resistances of conductors is given in the first section of this chapter.

For a *complete* circuit Ohm's law is equation (56),

$$\text{Current} = \frac{\text{electromotive force}}{\text{total resistance}}.$$

Electromotive force is equal to the total loss of potential in the circuit. Since the total loss must be equal to the total amount supplied by the cell or other generator, the electromotive force is also equal to the total amount that the potential is raised. It follows that the electromotive force is equal to the total electrical energy expended when a unit charge has flowed through the circuit.

If a part of a circuit contains not only a resistance but also an electromotive force, E, the difference of potential of the terminals of that part is given by equation (62),

$$V = E - Ir.$$

When there is a counter electromotive force, E, in this part of the circuit, the difference of potential is equation (63),

$$V = E + Ir.$$

Ohm's law does not apply to the conduction of electricity through gases and vapors but to the steady flow of current through metallic and liquid conductors. In dealing with alternating currents, one must use Ohm's law with caution, but it may be applied in the cases of incandescent lamps and electric heaters.

When electrical energy other than that used in developing heat is supplied to any part of a circuit, there is always present a counter electromotive force which opposes the flow of current. That part of the power which does not go into the heating of conductors is always equal to the product of the current and the counter electromotive force.

PROBLEMS

1. Two resistances, of 2 and 4 ohms, are connected in parallel. A resistance of 3 ohms is in series with the pair. Find the total resistance.

2. Five incandescent lamps, each having a resistance of 220 ohms, are joined in parallel across 110-volt mains. Find (*a*) the total current through the lamps, and (*b*) the total power in watts.

3. Two parallel resistances, of 5 and 10 ohms respectively, are connected in series with a resistance of 15 ohms. What is the current in each branch of the circuit when the total fall of potential is 10 volts?

4. Twenty similar lamps, in parallel, are supplied with 0.5 ampere each, with a difference of potential of 110 volts at the lamps. If 2.2 volts are lost on the line which connects them to the generator, what is (*a*) the voltage at the terminals of the generator? (*b*) the resistance of the line? (*c*) the power loss in the line?

5. A battery with an internal resistance of 2 ohms is connected, by wires with a resistance of 0.8 ohm, to an instrument which has a resistance of 2.2 ohms. If the current is 4 amperes, (*a*) find for each part of the circuit the loss of voltage produced by resistance. (*b*) What is the total loss? (*c*) What is the E.M.F. of the battery?

6. Two resistances, of 6 and 12 ohms respectively, are connected in parallel. To the combination is connected a cell with an E.M.F. of 1.5 volts and with an internal resistance of 1 ohm. (*a*) Compute the value of the current in the battery. (*b*) Find the current in each branch.

7. Two resistances, of 20 and 40 ohms, are connected in parallel. This group is in series with 10 ohms and with a battery the internal resistance of which is 5 ohms. If the E.M.F. of the battery is 10 volts, compute the current in the battery.

8. Two cells, each having an internal resistance of 3 ohms and an electromotive force of 1.5 volts, are connected in parallel. Find the total current when the circuit is completed through a resistance of 4 ohms.

9. A cell having a resistance of 2 ohms and an E.M.F. of 1.4 volts is connected to a coil with a resistance of 5 ohms. Compute (*a*) the current, and (*b*) the difference of potential between the terminals of the cell.

10. 5 amperes are to be supplied for house-lighting. The lamps, which are near together and in parallel, are connected to a storage battery by 250 ft. of wire which runs 1.6 ohms per 1000 ft. What must be the difference of potential of the battery terminals in order to have 30 volts across the lamps?

11. A dry cell with an E.M.F. of 1.5 volts and an internal resistance of 0.2 ohm is connected in series with a gravity cell having an E.M.F. of 1 volt and an internal resistance of 2 ohms. The circuit is completed by wires having a resistance of 2.8 ohms. (*a*) What is the current? (*b*) Compute the difference of potential at the terminals of each cell.

12. The connections of the circuit of Problem 11 are changed so that the two electromotive forces are opposing each other. What is (*a*) the current? (*b*) the difference of potential at the terminals of each cell?

13. A dynamo having an internal resistance of 0.2 ohm sends a current of 20 amperes through wires having a resistance of 1.3 ohms to a distant building, where the difference of potential of line is 105 volts. (*a*) What is the E. M. F. of the dynamo? (*b*) How much power is lost in the dynamo and the connecting wires?

14. A storage battery of a home-lighting plant has an E. M. F. of 32 volts and an internal resistance of 0.2 ohm. It is connected by wires having a resistance of 0.2 ohm to a group of 5 lamps which are in parallel. (*a*) If each lamp has a resistance of 30 ohms, what is the current supplied? (*b*) What is the difference of potential at the lamp terminals?

15. A current of 12 amperes divides, flowing through two parallel resistances of 1.5 ohms and 0.5 ohm. Find the current through the branch having a resistance of 0.5 ohm.

16. The moving coil of an ammeter has a resistance of 1.998 ohms, while the shunt has a resistance of .002 ohm. What is the current through each part when the total current is 20 amperes?

17. A certain ammeter has a resistance of 0.1 ohm. It is desired to add a shunt to it so that the ammeter proper and the shunt will together carry a current a hundred times greater than that indicated on the ammeter scale. What must be the resistance of the shunt?

18. Five lamps in parallel, each having a resistance of 30 ohms, are connected by wires with a resistance of 0.5 ohm to a dynamo which has a resistance of 0.2 ohm. What must be the E. M. F. in order that each lamp may have a current of 1 ampere?

19. A dynamo, the voltage across the terminals of which is 220 volts, is connected to a distant motor by wires having a resistance of 0.5 ohm. The loss of voltage in the wires is 20 volts. (*a*) How much current is supplied to the motor? (*b*) What is the difference of potential of the motor terminals?

20. The voltage across the terminals of a direct-current motor is 205 volts, and the current is 20 amperes. If the internal resistance is 0.5 ohm, what is the counter E. M. F. of the motor?

21. A dynamo having an internal resistance of 0.5 ohm is connected, by wires with a resistance of 1.5 ohms, to a place where the power is used. At this place a voltmeter shows 114 volts, and an ammeter 25 amperes. What is the E. M. F. of the dynamo?

22. A storage cell has a resistance of 0.1 ohm and an E. M. F. of 2.2 volts. (*a*) What resistance must be connected in series with the cell so that it can be charged on a 20-volt circuit with a current of 5 amperes? (*b*) What will be the difference of potential of the terminals of the cell during charging? (The E. M. F. of a storage cell is a counter E. M. F. when on charge.)

23. A dynamo, the internal resistance of which is 0.4 ohm and the E. M. F. 120 volts, is charging a series of 50 storage cells, each having an E. M. F. of 2.1 volts. The resistance of each cell is 0.01 ohm, and that of all connecting wires used is 0.6 ohm. What is (*a*) the current? (*b*) the difference of potential at the terminals of the dynamo? (*c*) the difference of potential at the terminals of the battery?

24. A storage battery is charged by a dynamo circuit with a current of 10 amperes. In order to get this current, a resistance of 3 ohms is in series with the battery. The resistance of the dynamo and the connecting wires is 1.5 ohms. A voltmeter connected to the battery terminals reads 60 volts. (*a*) How much power is supplied to the battery? (*b*) How much power is wasted in heat in the 3 ohms, the dynamo, and the connecting wires? (*c*) What is the E. M. F. of the dynamo?

25. The E. M. F. of a storage battery is 100 volts. When producing 10 amperes, the difference of potential at its terminals is 98 volts. What is its internal resistance?

CHAPTER XXXI

CONDUCTION OF ELECTRICITY THROUGH ELECTROLYTES

Introduction, 448. Electrolysis, 449. Historical note, 450. The electrolysis of water, 451. Electroplating, 452. Ions, 453. Faraday's laws of electrolysis, 454. A theoretical derivation of Faraday's laws; the absolute mass of a hydrogen atom, 455. The voltameter, or coulometer, 456. The theory of the simple cell, 457. Types of cells, 458. Storage cells, 459.

448. Introduction. In the conduction of electricity through metals there is no evidence of any transfer of the material of the conductor. Hence there is no reason for believing that the current is carried by atoms or molecules of these metallic conductors. It is now generally believed that current in solid conductors is carried solely by the migration of electrons—the very small, negatively charged particles which are far smaller than the atoms of any element.

But in the conduction of electricity through liquids the facts are quite different. Material is carried by the current. A familiar instance is that of electroplating: in the plating of silver, for example, silver atoms migrate through the liquid and are deposited on one of the plates which dips into the liquid. A great many illustrations could be cited, and some of these will be given later. The evidence shows that electricity is carried through most liquids in an entirely different way from that by which it is carried in solid conductors. The difference between the two is so great that the subject of the conduction of electricity through liquids is often treated as a separate branch of electricity.

449. Electrolysis. When a current flows through a liquid, not only is there a transport of material, but usually there are chemical changes in the liquid. For example, if a current flows through a solution of common salt (sodium chloride), entering and leaving the liquid by means of platinum terminals which dip into the liquid, sodium hydroxide, or caustic soda, is formed in the liquid near one terminal, and hydrochloric acid near the other. Liquids in which such changes are produced by electric currents are called *electrolytes*, and the whole process is called *electrolysis*. Faraday, who introduced these names, suggested also the names now commonly used for the terminals conducting the current into and from the liquid. These he called *electrodes*: the one which conducts the current into the liquid he called the *anode*, and the one which leads the current out, the *cathode*.

Mercury and other molten metals conduct in a manner similar to that of solid metals.

450. Historical note. On March 20, 1800, Volta wrote a letter to Joseph Banks, the president of the Royal Society of London, describing a method of making a battery with plates of zinc and copper and with pieces of cloth wet with a salt solution. These were laid one on top of another in the following order: zinc, cloth, copper, zinc, cloth, copper, zinc, cloth, copper, and so on. The zinc-cloth-copper forms a complete cell, and stacking up a large number of pieces forms a battery of many cells *in series*. Within six weeks after Volta's letter was received, a battery—the first in England—had been constructed; and by means of it the decomposition of water had been discovered by Nicholson and Carlisle. This discovery aroused great interest and led others to take up the work. Soon Cruickshank precipitated silver and copper from their solutions,—a discovery which led to electroplating. In 1807 Davy decomposed soda and potash, proving that they were not elements, as had been supposed, but chemical compounds. After this, progress was rapid.

Before the voltaic cell was discovered, there were no sources known that would give currents large enough to produce electrolytic effects which could be readily observed. Forms of the electrostatic machine were in use, but they were of little service in producing continuous currents. The discovery of the voltaic cell opened up a new era in electricity: the discoveries in electrolysis were only a part of the great discoveries that took place in the years from 1800 to 1830.

451. The electrolysis of water. The decomposition of water can be shown in a number of ways. Fig. 269 illustrates a form suitable for classroom projection. A cell of glass sides has platinum terminals (electrodes) running up through the bottom. The cell is filled with acidulated water to which a small amount of sulfuric acid has been added. Two glass tubes, closed at one end, are also filled with the liquid and are inverted over the electrodes. When a current flows through the liquid from one terminal to the other, gases form at the electrodes and, bubbling up, are caught by the inverted tubes. Oxygen, the gas formed at the anode, has about half the volume of the hydrogen formed at the cathode.

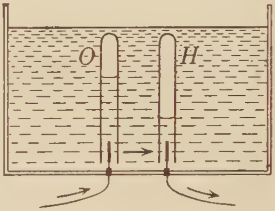


FIG. 269

452. Electroplating. When a current flows through a solution of copper sulfate, copper is *deposited* on the cathode and *dissolved* from the anode, if this electrode is copper. It is a general rule that the metal is deposited on the cathode. Hence the current must flow through the solution toward the object to be plated.

In silver-plating, a solution of potassium-silver cyanide is commonly used. It is important that a silver electrode be used for the anode; for if some other metal, say copper, were used, copper would be dissolved and the purity of the solution be spoiled.

Nickel-plating is also done by an electric current. In the case of most metals electroplating is very difficult, and it is often impossible to get firm, hard deposits. Copper, silver, and nickel are really exceptions. The "lead tree" is often shown to classes, to illustrate one kind of deposit which may be formed. A glass cell containing a solution of lead acetate and having lead electrodes is placed in a projection lantern. When a current is sent through the solution, instead of its depositing a uniform film over the cathode a mosslike growth of metallic lead forms on the cathode, developing in a lifelike manner and forming a beautiful pattern in the liquid.

453. Ions. Electricity is carried through a liquid by means of the migration of small particles of matter, each carrying a small charge. These carriers are called *ions*. The ions in electrolytes are atoms, and in some cases groups of atoms. Moreover, there are in a given electrolyte both positively and negatively charged ions; whereas in the conduction through metals there is only one kind of carrier—the negatively charged electron.

The ions in an electrolyte are formed from molecules which have been broken up or dissociated. The molecule of common salt (sodium chloride) is composed of two atoms: one of sodium and the other of chlorine. The sodium atom is charged positively; the chlorine atom, negatively. (It is generally believed that the attraction of the two unlike charges constitutes the bond which holds the two atoms together.) When the salt is dissolved in the water, the atoms may become separated, thus breaking up the molecule. However, the atoms retain their charges, so that there are scattered throughout the liquid many positively charged sodium atoms and an equal number of negatively charged chlorine atoms. When the electrodes dipping into the cell are connected to a battery and thus become charged, one positively and the other negatively, the charged atoms are attracted and repelled: the positively charged sodium atoms drift slowly toward the negative electrode, or cathode, and the negatively charged chlorine atoms drift toward the positive electrode, or anode. Thus the charged atoms become the carriers of the electric current. Occasionally two oppositely charged atoms may come close enough together to unite and form a molecule. But other molecules will be breaking up, so that the supply of ions is maintained. In dilute solutions practically all the molecules of the salt are broken up or dissociated.

According to the electron theory the sodium atom is positively charged because it has lost an electron, and the chlorine atom is negatively charged because it has picked up an extra electron. When ions such as these reach the electrodes, electrons are transferred between the electrodes and the ions. The positively charged ions, which go to the cathode, will receive electrons; and the negatively charged ones, which go to the anode, will give up electrons (this last statement must not be taken too literally, for often there are chemical changes at the electrodes which make the processes complicated).

In the case of the decomposition of acidulated water, the acid molecules are the ones that break up to form ions. The sulfuric-acid molecule consists of two hydrogen atoms, one sulfur atom, and four oxygen atoms. When it is dissociated, the hydrogen atoms become positive ions, and the cluster of one sulfur atom and four oxygen atoms forms a negative ion. The positive ions (the hydrogen atoms) go to the cathode, and hence hydrogen gas collects there. The negatively charged ions go to the anode (the positive terminal), where they unite with the molecules of water, forming sulfuric-acid molecules and releasing the excess of oxygen, which, in the form of a gas, bubbles off the anode.

The ions in electrolytes move with very slow speeds—approximately a centimeter per hour. The speeds of the different kinds of ions are different, even when they are acted on by the same electric forces. The slow speed of the ions does not mean that only very small currents can flow through an electrolyte. This speed is compensated by the large number of ions present—a number so great that a large quantity of electricity can easily be carried through in a short time.

454. Faraday's laws of electrolysis. The fundamental laws of electrolysis were first stated by Faraday. They may be stated in two laws.

1. *The mass of any substance deposited is proportional to the total quantity of electricity flowing through the solution; that is, to the product of the average current times the time.*

This law may be expressed in the form

$$\text{Mass deposited} = \epsilon Q, \qquad (67)$$

where Q is the quantity of electricity, and ϵ a constant which is called the *electrochemical equivalent*. If I is the average current and t the time, this equation may be written

$$\text{Mass deposited} = \epsilon It. \qquad (68)$$

A most important observation made by Faraday was that it does not make any difference in which part of a circuit the electrolysis takes place. For example, the same amount of copper is deposited if the electroplating apparatus is connected directly to one or the other terminal of the battery or placed in some more remote part of the circuit. This shows that the current is the same in all parts of a closed circuit.

2. *For the same quantity of electricity the mass of any element deposited is proportional to the chemical equivalent of that element* (the chemical equivalent of any element is its atomic weight divided by its valence).

According to this law, if a number of different electrolytic cells are connected in series so that the same current deposits different substances in the different cells, the masses deposited will not be the same. For example, the same quantity of electricity deposits over thirty times as many grams of copper and over a hundred times as many grams of silver as of hydrogen.

Another way of stating the second law is to say that the electrochemical equivalents of different elements are proportional to their chemical equivalents.

The following table gives the electrochemical equivalents in practical units of a number of elements; that is, the number to be used in equations (67) and (68) when coulombs or amperes are used.

TABLE OF ELECTROCHEMICAL EQUIVALENTS

(In Grams per Coulomb)

	Atomic Weight	Valence	ϵ
Copper	63.57	2	0.0003294
Hydrogen	1.008	1	0.00001046
Nickel	58.68	2	0.0003041
Nickel	58.68	3	0.0002027
Oxygen	16.00	2	0.00008291
Silver	107.88	1	0.00111800

455. A theoretical derivation of Faraday's laws; the absolute mass of a hydrogen atom. It has already been pointed out that a charge of electricity is always a whole multiple of a small charge which has been called the natural unit (sect. 391). This fundamental unit we shall call e. We shall assume that what the chemists call the valence of an element is merely the number of these unit charges that the atom carries with it.* Then the charge on each atom is ev, where v is the valence.

If N is the number of atoms deposited on the electrode, the quantity of electricity Q carried up to the electrode is given by

$$Q = Nev. \tag{69}$$

The total mass M carried up to the electrode is given by

$$M = Nm, \tag{70}$$

where m is the mass of the atom in grams, and N, as before, is the number of atoms carried to the electrode.

From equation (69), $N = \frac{Q}{ev}$.

Substituting this value of N in equation (70), we have

$$M = \frac{m}{ev} Q. \tag{71}$$

This equation shows that the mass deposited, M, is proportional to the quantity Q (Faraday's first law). It shows also that the mass deposited is proportional to the chemical equivalent, m/v (Faraday's second law).

Comparing equation (67) with equation (71), it is seen that the electrochemical equivalent is related to the atomic constants by the equation

$$\epsilon = \frac{m}{ev}. \tag{72}$$

One very interesting thing follows from this equation. The electrochemical equivalent ϵ has been determined very accurately for a number of elements of which the valence v is known. The numerical value of the fundamental charge e has been determined accurately by Millikan (sect. 538). It follows that the absolute mass, m, of these elements can be computed by the aid of equation (72). For example, it is thus found that *the mass of a hydrogen atom equals* 1.66×10^{-24} *gm.*

From this and the known mass of a cubic centimeter of hydrogen, *the number of atoms in 1 cubic centimeter of hydrogen at 0° C. and with a pressure of 76 centimeters equals* 5.41×10^{19}.

The chemists tell us that each molecule of hydrogen in the gaseous state is made of 2 atoms; hence *the number of hydrogen molecules in 1 cubic centimeter at 0° C. and with a pressure of 76 centimeters equals* 2.705×10^{19}.

---

* This follows at once from the theory that electrical attraction is what holds the atoms together. An atom of valence 2 has 2 unit charges; hence it requires 2 atoms which have only 1 charge of the opposite sign to form in combination with it a stable group. For example, in the case of the water molecule 2 atoms of hydrogen, each of which carries 1 positive unit, are required for each oxygen atom, which has 2 negative charges.

456. The voltameter, or coulometer. Since the electrochemical equivalents of many elements are now accurately known, the quantity of electricity that has passed through an electrolyte can be determined if the mass deposited can be accurately weighed. So precisely can this be done that it was decided by the International Electrical Congress of 1908 to define the ampere in terms of the rate at which it deposits silver. The *international ampere is defined as the unvarying current which, when passed through a solution of nitrate of silver* (prepared according to certain specifications) *deposits silver at the rate of .00111800 gram per second.*

Instruments designed for the purpose of measuring the quantity of electricity by measuring the deposit are called *voltameters* (not voltmeters). Since it is quantity of electricity that is measured, and since the unit for quantity is called the coulomb, the name *coulometers* is often used.

457. The theory of the simple cell. When a zinc plate is immersed in a dilute solution of sulfuric acid, there is a tendency for zinc atoms to leave the metal and go into solution. Some of them do. But each zinc atom that goes away carries with it a positive charge. The zinc plate will then have an excess of negative charge, and the solution will become positively charged. However, zinc atoms cannot continue indefinitely to leave; for the increasing negative charge on the zinc plate, together with the positive charge in the liquid, tends to prevent the positively charged zinc atoms from leaving the zinc plate. After the plate is put into the liquid, a condition of equilibrium is very quickly reached, when no more zinc atoms go into solution. Thus a certain definite difference of potential is produced between the liquid and the plate. This definite difference of potential is *maintained.* If for any reason the difference of potential becomes less, equilibrium is disturbed, and more positively charged zinc atoms leave the zinc; or, if the difference becomes too large, zinc atoms are forced back to the zinc plate.

When a copper plate is placed in acidulated water, there is a tendency for the positively charged hydrogen ions in the solution to leave the solution and go to the copper plate. This tends to raise the potential of the copper plate higher than the potential of the liquid, but the difference produced is not so great as that produced at the zinc-liquid surface. Fig. 270 shows diagrammatically the relative potentials when zinc and copper plates are placed in acidulated water. From this diagram, as well as from the explanation, it is seen that a definite difference of potential is produced between the copper and the zinc, the copper having a higher potential than the zinc.

What will happen when the copper and the zinc are united by an external metallic connection? It is apparent that a current will at once begin to flow from the copper to the zinc, and that this current flowing to the zinc plate will decrease the negative charge there. Just as soon as the negative charge on the zinc is decreased, the electrical equilibrium is disturbed, the attractive force holding the positively charged zinc atoms to the plate is decreased, and zinc atoms will go off into the solution. These zinc atoms, carrying into the liquid a positive charge from the zinc, tend to lower the potential of the zinc. Or, to make a long story short, the atoms will continue to leave the plate as long as the zinc and copper are connected through an external circuit. Current is thus maintained.

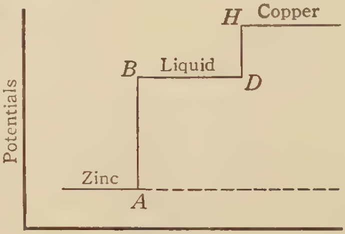


FIG. 270

But where is the source of the energy which supplies the work that is done when the current flows? When the zinc atoms go into solution, they lose energy; that is, their potential energy becomes less or, as it is usually stated, the chemical energy is less. Zinc in solution in dilute sulfuric acid has less potential energy than when in the form of the pure metal. It is just as proper to say that the zinc is "eaten up," or destroyed, as it is to say that coal is destroyed when burned. The carbon in the coal is not destroyed, but undergoes a chemical change with a loss of potential energy. The zinc in the cell changes its chemical form and loses energy. One reason why it is so expensive to supply any considerable amount of electrical energy by means of cells is that zinc is a very expensive fuel.

The diagram in Fig. 270 does not apply when the cell is producing current. In this case the liquid is not all at the same potential; for a liquid has resistance, and there is a fall of potential (current times resistance) in the liquid (Fig. 271). In Fig. 270 the vertical distance between the zinc and copper lines represents the electromotive force of the cell—the total amount the potential is raised (sect. 442). In Fig. 271 the vertical distance between the zinc and copper lines is the actual difference of potential of the terminals; it is equal to what a voltmeter would read when connected to the two. It is less than the electromotive force by an amount equal to the loss in potential in the liquid. That is, the difference of potential of the terminals is

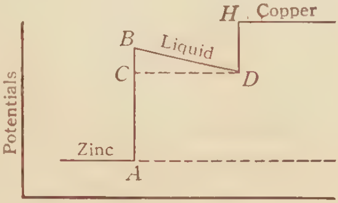


FIG. 271

$$\text{E.M.F.} - BC.$$

The electromotive force is, in both figures,

$$AB + DH.$$

458. Types of cells. In the *gravity* cell a copper plate is placed at the bottom, and the zinc "crowfoot" at the top. Copper-sulfate crystals are placed around the copper plate in the lower part of the cell. When the cell is in good condition, there is a fairly clear line of demarcation between the clear solution of zinc sulfate at the top and the blue copper sulfate solution at the bottom. This cell must be used continuously to be kept in good condition; hence it is well suited for giving a small, *continuous* current. The internal resistance is rather high—from 2 ohms up to 10 or more. The electromotive force is about 1 volt.

This cell illustrates well the principle of two-fluid cells. In a simple cell, such as one composed of copper and zinc immersed in dilute sulfuric acid, hydrogen will be deposited on the copper plate when the current is flowing. The copper plate soon becomes covered with a layer of small bubbles of the gas, which interferes with the action of the cell. This action is usually called *polarization*.

In the gravity cell the hydrogen, instead of forming on the copper plate, unites with the copper sulfate. The copper sulfate is called the *depolarizer*.

The *dry cell* is by far the most commonly used cell. It is well suited for intermittent service, such as doorbells and flash lights. The electromotive force is about 1.5 volts, and the internal resistance less than 0.1 ohm for the standard-sized cell when it is new. As the cell gets older the resistance increases, even though the cell may not have been used. An old cell may have a resistance of over 100 ohms. The cell is made of a zinc container, which is one of the terminals, with a packing of some absorbent material, such as pulpboard or blotting-paper, next to the zinc. A carbon plate or rod is placed in the center, and the space between this and the lining is filled with a mixture of granulated carbon and manganese dioxide, this last being the depolarizer. The absorbent layer is saturated with a solution of sal ammoniac and zinc chloride in water. The top is sealed with some pitch compound, to prevent the evaporation of the water. A good cell of the standard size, $2\frac{1}{2}$ by 6 inches, should give a discharge of about 25 ampere-hours (for example, 0.5 ampere for 50 hours, or 1 ampere for 25 hours) before being exhausted.

459. Storage cells. Accumulators, or, as they are more commonly called in this country, storage cells, have come into such extended use that a knowledge of some of their properties is very desirable. In the construction of the ordinary voltaic cell, plates of different materials are used (for example, they may be zinc and copper, or zinc and carbon), but in the storage cell the plates may be made of the same material. A simple storage cell can be made by immersing two plates of lead in dilute sulfuric acid. If a current from some source is sent through such a cell, chemical actions of different kinds occur at the two plates. The surface of one of these plates becomes changed to a peroxide of lead, and the other to spongy lead. The "active" substances of the two plates are now different, and this cell can be used to supply a current. If it does produce a current, this will flow through the cell in the *opposite* direction from that of the charging current. This discharging current will flow until the plates under the reversed chemical reactions become alike again. It will then be necessary to recharge the cell. The principle involved, stated in a very simple way, is that the charging current flowing through the cell makes the surfaces of the plates chemically different, so that chemical energy is stored on the plates. The two plates, being now different, will behave like an ordinary cell. It is chemical energy that is stored in the cell, and not electricity.

In order to facilitate the chemical actions that take place at the electrodes, none of the commercial forms of storage cells are built with plain electrodes. Usually a "pasted" plate is used. The plate, sometimes called the *grid*, is made with a large number of perforations, into which other substances are placed.

There are two types of cells in common use: one type consists of lead plates immersed in dilute sulfuric acid; the other, the alkaline type, is best represented by the Edison cell.

The lead type. When the lead cell is charged, the active deposit on the positive plate is lead peroxide. This plate has a reddish-brown color. The active substance of the negative plate is spongy lead. (The positive plate is the one of higher potential,—the one from which the current flows in the discharge.) When the cell is discharging, lead sulfate is formed on both plates.

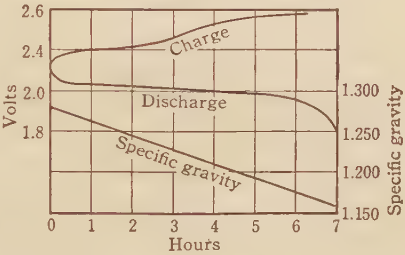


FIG. 272

When the cell is being charged, the density of the acid increases; when it is discharging, the density decreases. Therefore a test of the density of the acid affords a simple way of telling when the cell needs charging. The curves in Fig. 272 show the changes that take place in the voltage of the terminals on charge and discharge, and the changes in the density of the acid on discharge. The exact form of these curves depends on the size of the current in the charge or discharge, and on other factors.

Cells are rated in two ways. They often are rated by giving the "normal" current. This is the continuous current that the cell can generate for a certain number of hours; in stationary cells it is ordinarily 8 hours. Cells are rated also in ampere-hours. An 80-ampere-hour cell will give, approximately, 10 amperes for 8 hours, 5 amperes for 16 hours, 1 ampere for 80 hours, and so on.

The Edison cell. In this cell nickel-plated steel plates are immersed in a solution of potassium hydroxide. The Edison cell does not have as high an efficiency as the lead cell, and its electromotive force is not only lower but also less constant. Its average electromotive force on discharge is 1.2 volts, while that of the lead cell is about 2 volts. However, it can be completely discharged without injury, and may be left for some time in this condition. In general the Edison cell will stand rougher usage than the lead cell. The curves in Fig. 273 show the terminal voltage on charge and discharge.

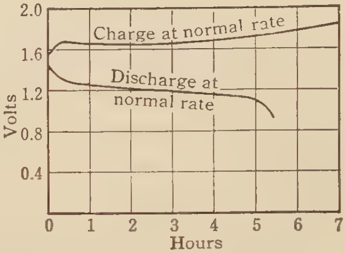


FIG. 273

PROBLEMS

1. The electrochemical equivalent of copper is 0.000329 gm. per coulomb. How many grams will be deposited by 1 ampere flowing for 1 hr.?

2. How many amperes are needed to deposit 3 gm. of silver in an hour? (The electrochemical equivalent of silver = 0.001118 gm. per coulomb.)

3. If 20 amperes are passed through acidulated water for 5 min., how many grams of hydrogen are liberated? (The electrochemical equivalent of hydrogen is 0.00001046.)

4. An electrolytic cell has a counter electromotive force of 2 volts and a resistance of 4 ohms. This is connected to a battery of 4 volts and an internal resistance of 1 ohm. (*a*) Find the current in the circuit. (*b*) Find the number of grams of hydrogen deposited in 20 min. if the electrochemical equivalent of hydrogen is 0.00001046.

5. The electrochemical equivalent of hydrogen is m/e (equation (72), sect. 455). According to Millikan, $e = 1.591 \times 10^{-19}$ coulombs. Compute the value of the mass of the hydrogen atom.

6. Taking the electrochemical equivalent of silver as 0.0011180 gm. per coulomb, compute the electrochemical equivalent of oxygen (use the atomic weights and valences given in section 454).

CHAPTER XXXII

MAGNETIC FIELDS OF CURRENTS; MECHANICAL FORCES ON CONDUCTORS IN MAGNETIC FIELDS

Introduction, 460. Quantitative relations between current and strength of magnetic field; definition of unit current, 461. The mechanical force acting on a current-carrying wire placed in a magnetic field, 462. Summary of facts about current-magnetic-field forces, 463. The force on a straight wire in a uniform magnetic field, 464. The attraction or repulsion of two parallel wires, 465. Coils of wire in magnetic fields, 466. The moving-coil galvanometer, 467. Direct-current ammeters, 468. Direct-current voltmeters, 469. Electric motors, 470. Summary, 471.

460. Introduction. It is worth while for the student to recall certain facts about magnetic fields produced by electric charges in motion, and about forces acting on coils of wire carrying a current when they are placed in magnetic fields. A number of these facts were stated in Chapter XXVII, but no full explanation was given of the relation between them. They may be summarized as follows:

1. When a current flows through a wire, a magnetic field is formed around the wire. Forces act on magnets brought near the wire.

2. The relation between the direction of the current and the direction of the magnetic field produced by the current is given by the right-hand rule (sect. 405). It is important that each student be familiar with this rule.

3. Currents flowing through coils, or helixes, of wire produce relatively strong fields inside these coils. When iron is placed inside a coil which has a current flowing through it, the iron becomes magnetized by induction (electromagnet).

4. If a current flows through a coil suspended in a magnetic field, there will be rotational forces (torques) acting on the coil.

The rotation of the coil of a galvanometer or of an ammeter, or the rotation of the armature of an electric motor, is an example of this action.

In Chapter XXVII experiments were described demonstrating the facts given above, but overlooking several important questions, among which are the following:

1. Why is there a force tending to move a current-carrying conductor when it is placed in a magnetic field?
2. Under what circumstances does this force exist?
3. What is the direction of this force? Does it bear any relation to the direction of the current and to the direction of the magnetic field?
4. On what does the magnitude of this force depend? How can it be made larger?

The student should be able to answer these questions before he has completed a reading of this chapter.

461. Quantitative relations between current and strength of magnetic field; definition of unit current. When a current flows through a wire, a magnetic field is produced. If the current is made twice as large, the strength of the field at any one place is doubled. In general, the strength of the magnetic field produced by a current is, at any one place, *proportional to that current.* This fact is utilized in the measurement of currents: in nearly all cases current is measured by means of the magnetic field it produces.

Not only is current *measured* by the strength of its magnetic field, but the *definition* of the *C.G.S. unit* of current is based on the strength of the field a current produces.

Unit current in the *C.G.S. electromagnetic* system may be defined as *the current that will produce at the center of a circular coil of wire of one turn a magnetic field the strength of which is* $\frac{2\pi}{r}$, where r is the radius of the circle.

If i units of current flow through the coil, the strength of the magnetic field, H, at the center is

$$H = \frac{2\pi i}{r}. \qquad (73)$$

If the coil has N turns, then

$$H=\frac{2\pi Ni}{r}. \tag{74}$$

The *ampere* is one tenth of the C.G.S. electromagnetic unit. Hence a current of 1 *ampere* will produce a field the strength of which is one tenth as great as that produced by the *C.G.S. electromagnetic* unit. If I represents the current in *amperes*, the strength of field at the center of a circular coil of N turns is

$$H=\frac{2\pi NI}{10r}. \tag{74 a}$$

This equation can be used either to compute the strength of the magnetic field produced at the center of a coil by a current I or, if the strength of the magnetic field is measured, to compute the value of the current.

Strength of magnetic field due to a long straight wire. The strength of the magnetic field near a long straight wire is proportional to the current, i, in the wire, and inversely proportional to the distance from the wire. Expressed in the form of an equation,

$$H=\frac{2i}{r}, \tag{75}$$

where H is the strength of the field, i the current measured in *C.G.S. electromagnetic* units, and r the distance of the point from the wire. Since 1 *ampere* is equal to one tenth of a *C.G.S. electromagnetic* unit, when the current is measured in amperes and indicated by I,

$$H=\frac{2I}{10r}. \tag{75 a}$$

Strength of magnetic field in a long solenoid. It can be proved that the magnetic field inside a long solenoid, or helix, is very uniform, and has a strength given by

$$H=\frac{4\pi iN}{l}, \tag{76}$$

where N is the total number of turns, and l is the length of the solenoid. If the current is measured in amperes,

$$H=\frac{4\pi IN}{10l}. \tag{76a}$$

Instead of defining unit current from equation (73), some prefer to use equation (75) or (76). The logical process is to choose some equation as the fundamental one and to define the unit of current from it. Then all the other equations which state the strength of the magnetic field produced by currents flowing in various-shaped circuits must be derived from the fundamental equation. But the methods are not simple, and no attempt is made here to show how any one of equations (73), (75), and (76) can be derived from any one of the others.

462. The mechanical force acting on a current-carrying wire placed in a magnetic field. In Fig. 274, *AB* represents a wire placed in the magnetic field of a magnet. When a current flows along this wire, there will be a force acting on the wire. This force will not be one of attraction nor of repulsion, but the wire will be pushed either toward the student or away from him, depending on the direction of the current. The direction of the *force* will be at *right angles to the wire* and also at *right angles to the magnetic field* produced by the magnet.

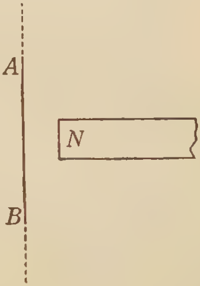


FIG. 274

The existence of this force is one of the most important facts of electricity and magnetism. In order to distinguish it from other forces that may be acting on the wire, it is called the *current-magnetic-field force.*

The current-magnetic-field force can be shown experimentally whenever a wire that is sufficiently flexible is used. The carbon filament of an incandescent lamp is usually flexible enough to show a deflection when a pole of a magnet is brought near. Another method is to suspend a loop of light tinsel wire; or, if ordinary wire is used, it can be suspended as shown in Fig. 275. The loop is suspended by means of hooks made by bending the ends of the bare wires, the hooks being turned in such a direction that they do not interfere with the swinging of the loop. Heavy wire should not be used.

Fig. 276 shows a sectional view of Fig. 275; that is, a section formed by a perpendicular plane cutting the wire. The section of the wire is at *P*. Let us suppose that the current in the wire of

Fig. 276 is flowing away from the student. Then the magnetic field will be in a clockwise direction around the wire, a few of the magnetic lines being shown in the figure (the student should verify the direction by the right-hand rule). It has been repeatedly pointed out that the direction of a magnetic field is the direction in which the north-seeking pole of a magnet is pushed. Hence, when a north-seeking pole is brought near this wire, there will be a force pushing the pole in the direction *B*, because the field produced at the pole by the current is in that direction. But mechanical forces are always exerted in pairs, and there must be, according to Newton's third law of motion, an equal but opposite force. One should expect, then, that there would be an equal and opposite force acting upon the wire which carries the current. This force, shown by the arrow *PM*, is the one we have called the current-magnetic-field force. It is now easily seen why this force must be perpendicular to the wire; for the magnetic field of the current, and hence the force *B*, which is in the direction of this magnetic field, is perpendicular to the current. Further, it should be clear that the force *PM* must also be perpendicular to the magnetic field produced by the pole.

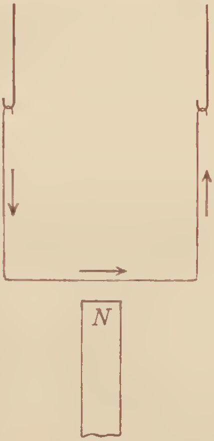


FIG. 275

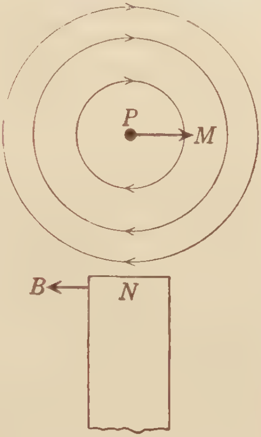


FIG. 276

A rule for finding the direction of the force on the wire can now be stated: *By the aid of the right-hand rule the direction of the force on the pole is found. The force on the wire is always opposite in direction to this.* Another rule, which in some respects is much better, is given in the following paragraph.

In Fig. 277 is shown part of the *resulting* magnetic field of the current and a magnet. On the right side of the wire the two fields oppose each other—the field due to the current being downward, while that due to the magnet is upward. The resulting field on this side is thus a weak one, with a small number of lines passing through it. On the left side of the wire the two fields are in the same direction, reënforcing each other and thus making a strong magnetic field with many lines of force. Hence lines crowd into this region and produce the distorted field shown in the figure. In section 348 we learned to regard these lines as being under tension and repelling each other sidewise. These tendencies cause the lines to push the wire to the side, as shown by the arrow *F*,—a direction perpendicular to the wire and to the field produced by the magnet. This method is very useful, for it gives one a simple, concrete mental picture of the process. Further than that, it shows clearly that the direction of the force on the wire depends (1) on the direction of the field in which the wire is placed and (2) on the direction of the current in the wire. As an illustration of this method, the student should imagine the case of an east-west wire carrying a current eastward through a magnetic field which is directed from south to north. By the aid of the right-hand rule one finds that above the wire the direction of the magnetic field due to the current is opposite to the external field, but below the wire the two fields are in the same direction. Hence the resulting field must be so distorted that the lines crowd in under the wire. The mechanical force acting on the wire will tend to raise the wire.

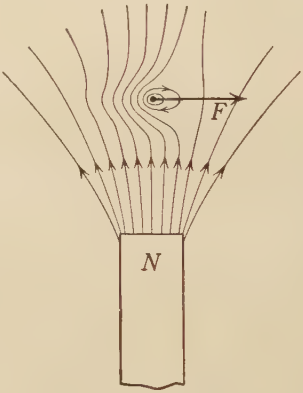


Fig. 277

With practice, one learns to simplify the mental process. For example, all that is necessary is to find out on which side of the wire the two magnetic fields are in the same direction; *the wire is always pushed away from the side where the field is strongest.*

State the direction of the force on the wire in the following cases:

1. If the current in Fig. 274 is flowing from A to B.

2. If the pole in Fig. 274 is changed to a south-seeking pole and if the current flows from B to A.

3. A vertical wire has a current flowing upward. Find the direction of the force on the wire due to the earth's magnetic field.

463. Summary of facts about current-magnetic-field forces. The following facts regarding the forces exerted by magnetic fields on wires carrying currents can either be proved experimentally by the apparatus shown in Fig. 275, or they may be deduced from the discussion of the last section.

1. The direction of the current-magnetic-field force acting on the wire is always perpendicular to the direction of the magnetic field produced by the magnet.

2. The direction of the force is always perpendicular to the direction of the length of the wire.

3. The greatest force is produced when the field of the magnet is perpendicular to the direction of the current, and the least (zero) when the field is parallel to the current.

4. Reversing the current, or reversing the direction of the magnetic field produced by the magnet, reverses the direction of the force.

In the usual light-circuit the current is *alternating*: one that goes first in one direction and then in the other. If such a current is used for the experiments that have been described, the direction of the force will reverse each time the current reverses, giving a force which reverses so many times a second that usually no effect is seen. A pole of a magnet brought near a carbon-filament incandescent lamp with an alternating current flowing through it will set the filament in very rapid oscillation instead of producing a deflection in one direction.

464. The force on a straight wire in a uniform magnetic field. The *magnitude*, as well as the direction, of the current-magnetic-field force is of interest. It is proved in more advanced texts that if a straight wire carrying a current is placed in and perpendicular to a uniform magnetic field the strength or intensity of which is H, the force acting on the wire is

$$F = Hil, \tag{77}$$

where i is the current measured in C.G.S. electromagnetic units, and l the length of the wire in *centimeters*. The force F is measured in *dynes*. If the current I is measured in *amperes*,

$$F=\frac{HIl}{10}. \tag{77a}$$

465. The attraction or repulsion of two parallel wires. In the case of two parallel wires carrying current, each wire produces a magnetic field; hence each wire lies in a magnetic field produced by the other wire. Let A and B in Fig. 278 be the sections of two wires perpendicular to the paper. If the current in B is flowing up toward the student, the magnetic field near A due to B will lie in the direction shown by the arrows. If there is a current in A, the force acting on A will be either to the right or to the left, depending on the direction of the current in A. Hence the force on A will be either an attraction or a repulsion. The student should have little trouble in completing the formulation of the following rules:

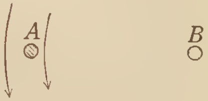


FIG. 278

1. *If the two currents are flowing in the same direction, the force will be one of* ——.

2. *If the currents are flowing in opposite directions, the force will be one of* ——.

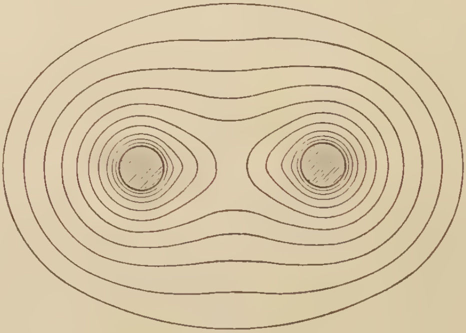


FIG. 279

In Fig. 279 is shown the magnetic field when the currents are flowing in the same direction, while Fig. 280 shows it for the case where the currents are in the opposite direction. In both cases the figures give sections taken perpendicularly to the wires.

466. Coils of wire in magnetic fields. In section 406 experiments were described showing the tendency of a current-carrying coil to rotate when placed in a magnetic field. This action will now be explained. In Fig. 281, *AB* is the coil, shown in cross section. If the coil is suspended as in a galvanometer, the section is a horizontal one. The magnetic field is produced by the two poles. By applying either of the two methods given in section 462, the student will see that the arrows give the direction of the current-magnetic-field forces for the case where the current of *A* is flowing away from him, and the current at *B* is flowing toward him.* As the coil moves, the forces stay in the indicated direction, and tend to rotate the coil until it has turned through a right angle; that is, until its plane is perpendicular to the magnetic lines between the poles.

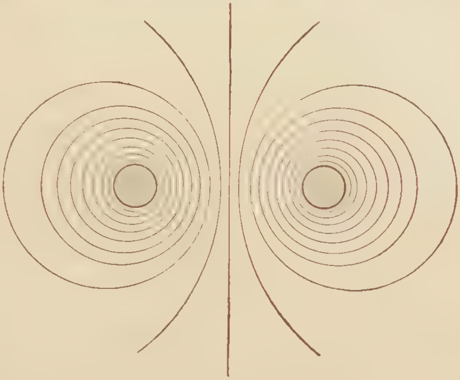


FIG. 280

In Fig. 282 is shown part of the *resulting* magnetic field. The distortion shows that there will be a torque, tending to rotate the coil in a counterclockwise direction.

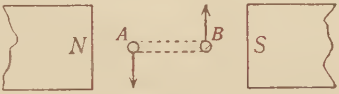


FIG. 281

There is another way of explaining the action of coils which carry currents. A coil of wire in which a current flows has a magnetic field similar to that produced by a magnet of the same size and shape as the coil. In other words, a coil of wire with a current flowing through it behaves as if it were a magnet. In Fig. 283 is shown the section of a coil formed by three turns of wire. If the current is flowing away from

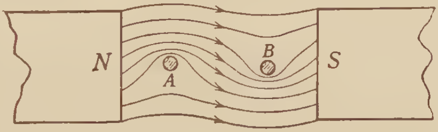


FIG. 282

---

* The direction of the force on *A* or *B* is the same whether the *N* pole or the *S* pole, or both, are used; for the direction of the force depends on the direction of the magnetic field, and the two poles produce fields in the same direction at both *A* and *B*.

the student on the upper side and toward him on the lower side, the direction of the magnetic field will be as shown. This coil has the same sort of magnetic field and produces the same forces as would be produced by a short thick magnet (shown, in section, by the dotted lines). The north-seeking pole is on the left face, and the south-seeking pole on the other face, as shown.

The advantage of regarding the coil as equivalent to a magnet is that one can often more quickly determine the character of the forces when the coil is placed in a magnetic field. For example, consider the case illustrated in Fig. 284. In this figure the current is flowing in at *A* and out at *B*; hence the coil may be regarded as equivalent to a short magnet, with poles which are marked *N′* and *S′* in the figure. From the known action of poles on one another the rotational forces will be in the direction indicated by the arrows. Compare this case with Fig. 281, where the current and the field have relatively the same direction. In both cases it should be clear that the coils tend to rotate through a quarter-turn, or until the magnetic fields of the coil and the magnet are in the same direction.

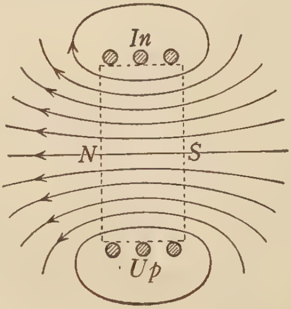


FIG. 283

It can be stated as a general rule that *a coil tends to turn into the position where the number of magnetic lines passing through the coil will be a maximum.*

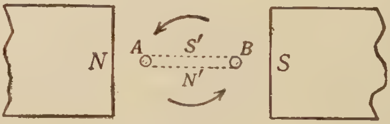


FIG. 284

The student should now be able to answer each of the questions in section 460.

467. The moving-coil galvanometer. Galvanometers are commonly used in the laboratory to measure or to detect small currents. In many cases it is desirable that they be made very sensitive. It is not difficult to understand some essential points which affect the sensibility.

When the suspended coil (Fig. 285) rotates on account of the torque due to the current-magnetic-field forces, it turns against the torque of the fine wire by which the coil is suspended. The smaller the wire used, the farther will the coil turn. To make the opposing torque of this wire very small, the wire is rolled out into a thin, flat, narrow ribbon. In this form it has approximately the same tensile strength, which it needs to support the coil, but it exerts a smaller torque. The lower connecting wire is usually coiled into a spiral, thus increasing its length and diminishing its torque.

The stronger the magnetic field in which the coil is placed, the greater the torque tending to turn the coil when a current flows through it. The magnets are hence made of special steel which will produce strong fields. The field is often made stronger and more uniform by supporting from the frame of the instrument a soft-iron core in the center of the field (*C* in the figure).

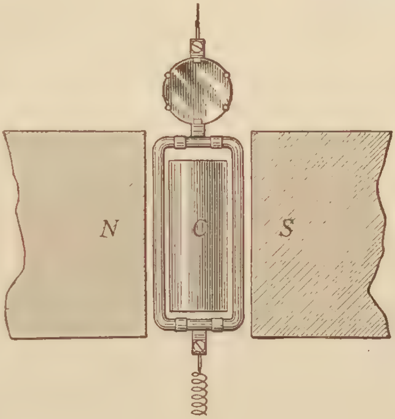


FIG. 285

The greater the number of turns of wire on the coil, the greater the resultant torque; for there is a torque acting on *each* turn. But winding on a large number of turns requires the use of small wire. Decreasing the size of the wire and increasing its length increases the resistance of the coil. An increased resistance is for some kinds of work almost prohibitive. Consequently galvanometers of both kinds are made: some with coils with many turns of fine wire, having a high resistance; others with a smaller number of turns, of a larger wire, having a low resistance.

468. Direct-current ammeters. The more common type of direct-current ammeters has a coil mounted between the poles of a strong permanent magnet, as in the case of a moving-coil galvanometer. These coils, instead of being suspended with long thin wires, are mounted in jewel bearings, and held in place by two flat spiral springs similar to the hairspring of the balance wheel of a watch.

Since the moving coil of an ammeter is entirely too delicate and sensitive to permit large currents to flow through it, most of the current flows through a shunt across the terminals. This arrangement is shown in Fig. 286. *C* represents the moving coil, and *S* the shunt. If 0.01 ampere flowing through the coil will give the maximum deflection, and if it is desired to have an instrument that will read a maximum of 10 amperes, the resistance of the shunt must be of such value that when the total current is 10 amperes, 9.99 amperes will flow through the shunt and 0.01 ampere through the coil. For other values of the current, the part flowing through the coil will always be proportional to the total current (see equation 54, sect. 440). It is customary to number the scale so that it indicates the total current, and not the current through the coil.

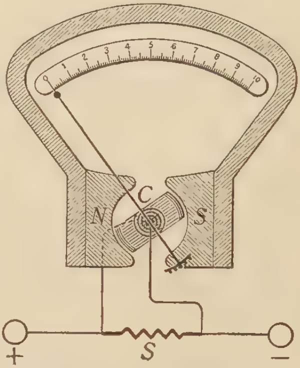


FIG. 286. An ammeter

469. Direct-current voltmeters. Although the electrostatic voltmeter is commonly used for certain types of work, especially for high voltages, the moving-coil types are frequently used in laboratories when the voltage is that of cells or power circuits which give direct or continuous currents.

The moving coil of a voltmeter is similar to that of an ammeter. There is no shunt, but a high resistance is inserted in series with the moving coil. A voltmeter is connected to the circuit in a different manner from an ammeter. The ammeter is always in series, so that all the current flows through it; the voltmeter is connected across the circuit to the two lines the *difference* in potential of which is to be measured. Fig. 287 shows symbolically how an ammeter and a voltmeter are connected to the two wires of a battery circuit. The shunt of the ammeter and the series

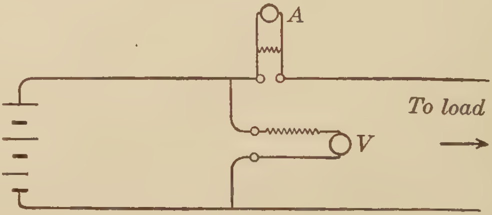


FIG. 287

resistance of the voltmeter are shown in the figure. Usually these are on the inside of the instruments and not visible.

It is always a matter of some interest and curiosity to know why two such similar instruments as an ammeter and a voltmeter measure two different things. To aid the student an outline of the theory of the voltmeter will be given here.

The deflection of a voltmeter is proportional to the current which flows through its moving coil. This can be stated by

$$I = kd, \tag{78}$$

where I is the current through the voltmeter coil, k is a constant, and d is the deflection. By Ohm's law the potential difference which is to be measured by the voltmeter is equal to the product of the current through the voltmeter, I, and the resistance of the voltmeter, r, or

$$V = Ir.$$

If the value of I from equation (78) is substituted in this equation, we obtain

$$V = kdr. \tag{79}$$

Both k and r are constants; hence equation (79) shows that the deflection, d, is proportional to the potential difference, V, of the points to which the voltmeter is connected. Since the scale may be marked in any desired manner, it can be graduated to indicate volts.

470. Electric motors. The tendency of coils placed in magnetic fields to rotate when the current flows through them is used in many types of electric motors. A simplified type of direct-current motor will be explained in this section.

In Fig. 288 is shown diagrammatically a coil mounted in a magnetic field on an axis AA', about which it can rotate. If the connecting wires were permanently fastened to the terminals of the coil, and a current sent through the coil, it would rotate only part way round. But if the terminals of the coil slide over the fixed ends of the connecting circuit in such a way that, as the coil turns around, each terminal of the coil slides off one of the

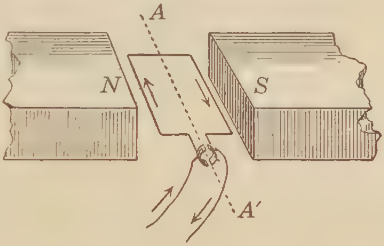


FIG. 288

fixed terminals and, a little farther on, slides onto the other fixed terminal, a continuous rotation can be produced. In Fig. 288 the terminals of the coil are shown as curved strips (see also Fig. 289), which, as the coil rotates, slide past the fixed terminals, or "brushes." With this arrangement the current in the side of the coil which happens to be next to the north-seeking pole will always flow away from the student, with the result that the current-magnetic-field force on that side of the coil is always acting downward (the student should verify this). The rotating part of the arrangement by which the current is changed from one wire to another is called the *commutator*, and the strips to which the ends of the revolving coil are fastened are called the *segments*.

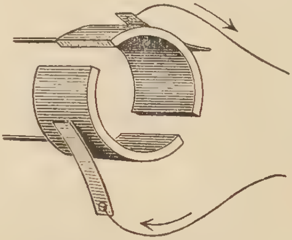


FIG. 289. A simple commutator

When the coil has turned through an angle of 90 degrees from the position shown in the figure, there will be in that position no torque tending to rotate the coil. Hence the torque will be neither continuous nor very effective. To overcome this a large number of coils are placed on the rotating part, with their planes inclined to each other, and instead of there being two segments in the commutator, a large number—thirty or forty—are used.

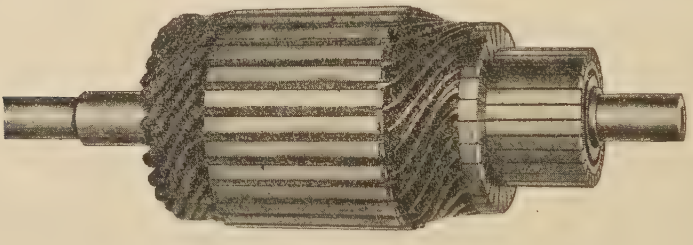


FIG. 290. The armature of an automobile starting motor

In Fig. 289 is shown a simple commutator. Fig. 290 shows a commutator (at the right end) and the coils of one type of commercial armature.

The series motor. The moving coil, or *armature*, of a motor is not placed between the poles of a permanent magnet but between those of an electromagnet which is magnetized by a current from the same source of supply that gives the current to the rotating coil. The advantage is obvious, for an electromagnet gives a much stronger magnetic field. There are several ways of connecting the rotating coil and the magnet coils. The simplest of these is shown diagrammatically in Fig. 291. The magnet coils *A* and *B*, called the *field windings*, are in series with the revolving coil *C*, or armature. This type is called a *series* motor. (The armature that is shown in Fig. 290 is part of a series motor.)

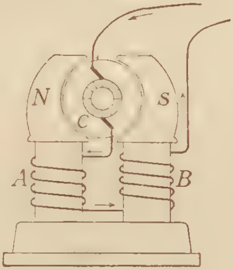


Fig. 291. A series motor

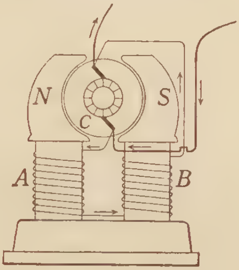


Fig. 292. A shunt motor

The shunt motor. In the *shunt* motor the armature *C* is connected in parallel with the field coils. As shown in Fig. 292, the current flows in on one wire to one brush. There it divides, part going through the armature *C*, the other part going through the field coils. These two currents unite again at the other brush. In this type of motor the field magnets are wound with many turns of fine wire.

The compound motor. The *compound* motor is a combination of the series and shunt motors. There are two windings on each pole. One of these is of fine wire in many turns, and is connected like that of the shunt motor. The other winding consists of a few turns of large wire which is connected in series.

Each of these types of motor has its own special advantages. The series motor will run at a high speed when it has a light load and at a low speed with a heavy load. The shunt motor has a fairly constant speed for different loads. A full account of the advantages and disadvantages is obviously beyond the scope of this book.

471. Summary. If the current is flowing in a circular coil of radius r and N turns, the strength of field H at the center is given by equation (74),

$$H = \frac{2\,\pi N i}{r},$$

where the current, i, is measured in C.G.S. electromagnetic units.

If the current is measured in *amperes* and denoted by I, then (equation (74 *a*))

$$H = \frac{2\,\pi N I}{10\,r}.$$

At a distance r from a long straight wire carrying a current i, measured in C.G.S. electromagnetic units (equation (75)),

$$H = \frac{2\,i}{r};$$

or, when current is measured in amperes (equation (75 *a*)),

$$H = \frac{2\,I}{10\,r}.$$

At the center of a long solenoid (equation (76)),

$$H = \frac{4\,\pi i N}{l},$$

where i is the current in C.G.S. electromagnetic units, N the total number of turns, and l the length of the solenoid.

The C.G.S. electromagnetic unit of current is defined in terms of the strength of the magnetic field it will produce. In the text, equation (74) is chosen for defining this unit.

Because a current flowing in a wire produces a magnetic field, it will exert mechanical forces on magnets close by. Since the current exerts a force on the magnet, the latter must react with a force which shows itself as a force on the wire which carries the current. Hence it follows that there will be forces acting on current-carrying conductors which are placed in magnetic fields. These forces are called *current-magnetic-field forces.*

The current-magnetic-field force is always in a direction perpendicular to the wire and also perpendicular to the magnetic field in which the wire is placed. Two methods are suggested in the text for finding the direction of this force.

When a straight wire carrying a current, i, measured in C.G.S. electromagnetic units, is placed in and at right angles to a uniform magnetic field, the current-magnetic-field force on the wire is (equation (77))

$$F = Hil,$$

where H is the strength of the field, and l is the length of the wire.

When two parallel wires are near each other, each wire produces a magnetic field near the other wire, so that each wire is in a magnetic field. These fields will be perpendicular to the wires, and a force, given by equation (77), will act on each wire. If the currents are flowing in the same direction, there is attraction; if they are flowing in opposite directions, there is repulsion.

A coil of wire carrying a current and placed in a magnetic field tends to set itself in such a position that the maximum number of magnetic lines thread through the coil.

Explanations of moving-coil galvanometers, direct-current ammeters, and voltmeters are given in this chapter.

Electric motors apply the forces developed when wires carrying currents are placed in strong fields.

PROBLEMS

1. A magnetic pole of strength 40 C.G.S. units is placed at the center of a circular coil which has a radius of 30 cm. and 10 turns of wire. Find the force on the pole when the current is 5 amperes.

2. A magnetic pole of 20 C.G.S. units is placed 5 cm. away from a long straight wire. What will be the force on this pole when the current in the wire is 20 amperes?

3. A straight wire carrying a current of 20 amperes runs perpendicularly through a magnetic field the strength of which is 5000 C.G.S. units. How much force is acting on each centimeter of length of the wire?

4. A north-seeking pole is placed just below an east-and-west wire carrying a current eastward. What is (*a*) the direction of the current-magnetic-field force on the wire? (*b*) the force on the pole? (Neglect the effect of the earth's magnetic field.)

5. An east-and-west wire carries a current of 10 amperes westward. (*a*) If the strength of the earth's field is 0.6 C.G.S. unit, what is the current-magnetic-field force on 1 m. of this wire? (*b*) If the angle of dip is 70°, what is the direction of this force?

6. A rectangular coil, 10 by 20 cm., of 15 turns, is hung in the earth's magnetic field with the plane of the coil parallel to the earth's field and in a vertical position. The current, 5 amperes, flows down on the north side of the coil and up on the south side. The horizontal component of the earth's field is 0.18 C.G.S. unit. (*a*) What is the direction of the current-magnetic-field force on the north edge of the coil? (*b*) Compute the torque acting about a vertical axis on the coil.

7. A long vertical wire carries a current of 25 amperes directed upward. A north-seeking pole of strength 120 C.G.S. units is placed 5 cm. north of the wire. (*a*) What will be the direction of the current-magnetic-field force? (Neglect the effects of the earth's field.) (*b*) What will be the magnitude of the force on the wire? (Find the magnitude of the force on the pole, as in Problem 2.)

8. An ammeter has a moving coil which, including connecting wires, has a resistance of 8.0 ohms. A current of 0.01 ampere through this coil will give a full-scale deflection. What must be the resistance of a shunt in order that a full-scale deflection may be produced by a total current of 10 amperes?

9. If the moving coil has a resistance of 5 ohms and gives a full-scale deflection for 0.011 ampere, how much resistance must be placed in series with it to make a voltmeter which will give a full-scale deflection with 3 volts?

10. What must be the value of the series resistance in order to make the instrument of Problem 9 a voltmeter which gives a full-scale deflection with 100 volts?

11. A voltmeter having a range of 0–5 volts has a total resistance of 450 ohms. What must be done to make the range 0–50 volts?

12. Two accurate voltmeters which have resistances of 10,000 and 15,000 ohms respectively are connected in series across a direct-current power circuit of 100 volts. What will the two instruments read?

CHAPTER XXXIII

INDUCED ELECTROMOTIVE FORCES

Introduction, 472. Faraday's experiment, 473. Other experiments, 474. General case, 475. A quantitative law of induced electromotive forces, 476. Induced electromotive forces produced by means of alternating currents, 477. The transformer, 478. A wire cutting magnetic lines, 479. Rules for the direction of the induced electromotive force, 480. Lenz's law, 481. Relationship between the current-magnetic-field force and the direction of the induced electromotive force, 482. The Arago experiment, 483. The induction coil, 484. The telephone transmitter, 485. A simple telephone circuit, 486. Self-induction, 487. Inductance, or coefficient of self-induction (L), 488. Energy in a magnetic field, 489. Dynamos, 490. The counter electromotive force of a motor, 491. The alternating-current dynamo, 492. Rotating magnetic fields: the induction motor, 493. Long-distance transmission of energy, 494. Summary, 495.

472. Introduction. The discovery of the voltaic cell opened a new era in electricity, for before that time only electric currents of insignificant magnitude could be produced. But even the voltaic cell is not suited for developing large amounts of electrical energy. The discoveries of Faraday and Henry opened up better methods for the production of electrical energy in great quantities. In fact, tremendous industrial applications of electricity today are possible only because of the discoveries of these men.

This chapter explains the principles involved in these discoveries, and their applications in such things as the dynamo, the transformer, and the telephone.

473. Faraday's experiment. Faraday wound two coils on an iron ring (see Fig. 293). These two windings were insulated from each other so that the current from one coil could not flow over into the other. Faraday found that when a current was starting in the coil connected to the battery, there was a *momentary* deflection of the galvanometer, G. He found also that when the current was broken by opening the key, K, there was another *momentary* deflection of the galvanometer, but opposite in direction to that obtained when the circuit was closed.

Faraday's experiment can be better understood if the student remembers that the iron ring becomes magnetized when the current flows through the coil, and loses its magnetism (or, at least, part of it) when the key is opened. The experiment showed that increasing the magnetic field through a coil connected to a galvanometer produced a momentary current in one direction, and decreasing the magnetic field in the coil produced a momentary current in the opposite direction.

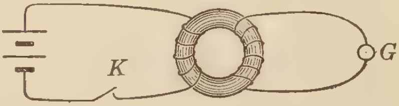


FIG. 293

474. Other experiments. When a bar magnet is moved in and out of a coil of wire connected to a moderately sensitive galvanometer, the following results can be observed (no battery is used):

1. When the magnet is moved into the coil, a deflection of the galvanometer is obtained.
2. When the magnet is taken out of the coil, a deflection in the opposite direction is obtained.
3. Current flows *only while the magnet is in motion.*
4. Moving the magnet rapidly produces a larger current than when it is moved slowly.

If another coil, connected to a battery, is brought near the coil which is connected to the galvanometer, the following results can be observed (the coil connected to the battery is called the *primary*, and the one connected to the galvanometer the *secondary*):

5. Closing the circuit in the primary produces a momentary current in the secondary.
6. Breaking the primary circuit produces a momentary current in the secondary opposite in direction to that of 5.
7. While current is flowing in the primary, moving the primary nearer the secondary produces a temporary current in the secondary.
8. Taking the primary away while current is flowing in it produces a temporary current opposite in direction to that of 7.

475. General case. Each of the experiments of the last section is a special case of the following:

An induced current is produced in a coil while the magnetic field through the coil is being increased or decreased.

Increasing the magnetic field produces a current in one direction, and decreasing the field produces a current in the opposite direction.[*]

476. A quantitative law of induced electromotive forces. In the preceding sections it has been stated that a *current* is produced while a magnetic field threading or interlinking a coil is changing. It has been found by simple measurements that what is really developed is an *electromotive force*, and that the current is a secondary effect. The law is as follows: *In any coil of wire the induced electromotive force is equal to the rate of change of the number of magnetic lines*[†] *in the coil multiplied by the number of turns of wire in the coil.*

The student should see in this rule two different factors on which the magnitude of the induced electromotive force depends:

1. The *time rate* of change in the number of magnetic lines; that is, the change per second in the number of lines.

2. The number of turns of wire in the coil. An electromotive force is induced in each turn of wire; and since these turns are in series, the total induced electromotive force is the sum of those in all the turns.

The law can be stated in the form of an equation:

$$e = N\frac{\phi_2 - \phi_1}{t}, \tag{80}$$

where e is the induced electromotive force, ϕ_1 the number of magnetic lines threading through the coil at the beginning, ϕ_2 the number at the end, t the time of the change in seconds, and N the number of turns of wire in the coil.

---

[*] In 1830–1833 Michael Faraday, at the Royal Institution, London, and Joseph Henry, at Albany Academy, Albany, New York, independently determined by experiments the circumstances under which an induced current could be produced. Although Faraday is generally given credit for the discovery because he published his results first, some phases of the work are clearly due to Henry.

[†] See section 365.

This rule gives the value of the electromotive force in the C.G.S. electromagnetic system of units. If the answer is desired in volts, the product given above must be divided by 100,000,000, since

$$1 \text{ volt} = 100{,}000{,}000, \text{ or } 10^8, \text{ C.G.S. electromagnetic units.}$$

Example. Take the case of a coil having 100 turns and an area of 500 sq. cm. (diameter about 10 in.). Suppose that the coil is held in a plane perpendicular to the earth's magnetic field and then, in one tenth of a second, is rotated through 90° about an axis in the plane of the coil. All magnetic lines are cut out in one tenth of a second. If the strength of the earth's magnetic field is 0.6 C.G.S. unit, the total number of magnetic lines threading through the coil is

$$\text{Strength} \times \text{area} = 0.6 \times 500 = 300 \text{ lines.}$$

If these lines are cut out in one tenth of a second, the *rate* of cutting is 300 divided by 1/10, or 3000 lines per second. The total E.M.F. is equal to the rate of cutting times the number of turns, or

$$3000 \times 100 = 300{,}000 \text{ electromagnetic units.}$$

Expressed in volts, $$\frac{300{,}000}{100{,}000{,}000} = 0.003 \text{ volt.}$$

This is a small electromotive force, but it is sufficiently large to produce a current which will give an appreciable deflection of a moderately sensitive galvanometer connected in series with the coil.

477. Induced electromotive forces produced by means of alternating currents. In most cases in America the current used for lighting houses and buildings is an alternating current; that is, the current flows first in one direction, then in the other, then back in the first direction, and so on. It does this rapidly, usually making about sixty complete "cycles" per second.

If an alternating current is available, some striking experiments with induced electromotive forces can be shown. In a coil of 150 or 200 turns of wire is placed a straight iron core 50 or 80 square centimeters in cross section, and 25 or 30 centimeters long, best made of strips of sheet iron or of pieces of iron wire. When the coil is connected to the alternating-current mains, the current will magnetize the iron core, but the magnetism will be continually changing. At one instant one end will be a north-seeking pole, a little later a south-seeking pole, then a north-seeking pole, and so on. If a 60-cycle current is used, the upper end will be a north-seeking pole 60 times a second, changing from a north-seeking pole to a south-seeking pole in $\frac{1}{120}$ second. With such a rapid rate of change of a magnetic field, relatively large electromotive forces can be induced. If another coil, a secondary, is placed near this coil, it is possible to light an incandescent lamp which is connected to the secondary.

A sheet of copper gets hot when held near the alternating-current magnet. Why? A frying-pan held against one pole may get hot enough to fry eggs.

If five or six turns of very heavy copper wire are wound in a coil, with its ends connected by a short piece of small copper wire, the small copper wire may be melted by the induced current when the coil is slipped down over the end of the alternating-current magnet.

478. The transformer. When Faraday wound an iron ring with two sets of windings,—a primary and a secondary,—he really constructed what we now term a *transformer*. Transformers are extensively used in connection with the distribution of electrical energy. For example, in nearly all cases where houses or buildings are lighted from a city circuit, the current flowing through the lamps does not come from the distant power house but from the secondary winding of a transformer. In the residential part of a city these transformers, in iron boxes, are mounted on poles in the streets or alleys.

The transformer is not a complicated machine, and its elementary principles are easily understood. The primary and secondary coils are usually placed close together, with an iron core extending through them. Usually the iron extends not only through the coils but around them, so that the magnetic lines are in iron throughout their entire path (see Figs. 294 and 295). When an alternating current is sent through the primary coil, the iron is magnetized first in one direction and then in the other, the magnetic field changing in intensity and direction with the same frequency as the current. This changing magnetic field induces an alternating electromotive force in the secondary, of the same frequency as the current in the primary. The magnitude of the induced electromotive force in the secondary will depend not only on the rate of change of the magnetic field through it but also on the number of turns in the secondary. Hence, by a proper adjustment of the number of turns in the secondary, an induced electromotive force of any desired magnitude can be obtained. When the number of turns in the secondary is smaller than that of the primary, the induced voltage is smaller than that applied to the primary. Such a transformer is called a *step-down* transformer. This is the kind commonly used in lighting-circuits. However, when the secondary contains a larger number of turns than the primary, the induced electromotive force in the secondary is correspondingly larger than that in the primary. Such a transformer is called a *step-up* transformer.

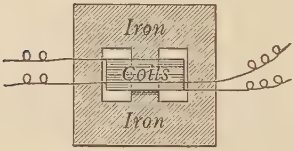


FIG. 294. A shell-type transformer

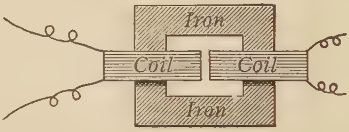


FIG. 295. A core-type transformer

When a large amount of power is transmitted at a low voltage, the current must be large, for power is proportional to the product of voltage and current. But a large current produces large heat losses along the transmission line. The remedy is to transmit the power at a high voltage and small amperage and then, by means of transformers, to step the voltage down to some convenient and safe value at each place where the power is used. A common voltage for transmission in a city lighting system is 2200 volts. Since this is entirely too large a voltage to be used inside a building, the wires from the mains are connected to the primary coils, or (as they are more commonly called) the *high-tension* coils, of the transformer. These coils have a relatively large number of turns. The wires going into the house are connected to the secondary, or *low-tension*, coils, which have a relatively small number of turns and hence a small induced voltage—usually 110 volts.

Small step-down transformers are commonly used in houses for supplying power to ring doorbells. The high-tension side is usually connected directly to the light-circuit, and the low-tension side, of a few volts, is connected to the bell circuit.

479. A wire cutting magnetic lines. There is another way of considering the mode of producing induced electromotive forces. So far we have regarded these electromotive forces as produced by changes in the magnetic field which threads through the coils. There are good reasons for thinking that a more fundamental view is to assume that the induced electromotive force is produced whenever magnetic lines move across wires, or wires move across the magnetic lines. Consider the rectangular coil *ABCD* placed in a magnetic field which has the direction shown in Fig. 296. Suppose this coil to be rotated about the side *AD* as the axis, the side *BC* being brought out toward the student. This motion increases the number of magnetic lines through the coil. During this motion the edges *AD*, *AB*, and *DC* do not cut magnetic lines. The side *BC* is the only one which cuts across the field. Hence, if we follow the point of view that an induced electromotive force is always produced by wires cutting across a magnetic field, we should consider the induced electromotive force as in the side *BC*.

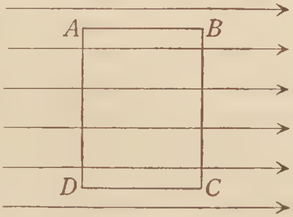


FIG. 296

In section 476 a rule was given for computing the magnitude of the induced electromotive force in terms of the number of lines threading through the coil. A rule stated in terms of the cutting of magnetic lines is as follows: *The magnitude of the induced electromotive force produced by a wire cutting across a magnetic field is equal to the number of lines cut per second by the wire.* This rule gives the value of the induced electromotive force in C.G.S. electromagnetic units.

This method and the one given in section 476 are not so different as appears at first thought. Magnetic lines have no free ends, and it is impossible to increase or to decrease the number threading through a closed coil *unless these lines cut across the wires of the* *coil.* Hence, whenever the number of lines in a coil is changing, the lines are cutting conductors. By using a simple case (that of Fig. 296, for example) the student should be able to prove that this rule and that of section 476 give identical numerical results.

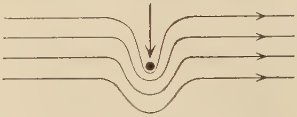


FIG. 297

480. Rules for the direction of the induced electromotive force. *Case 1: a wire moving across a magnetic field.* A very simple rule is obtained by making a simple assumption; namely, that when a wire moves across a magnetic field, or the field moves past the wire, the wire distorts the field as shown in Fig. 297. In this figure the wire is perpendicular to the plane of the paper and moves downward, as shown by the arrow.

If the wire is grasped with the right hand, with the fingers pointing in the way the distorted lines go around the wire, the thumb, which lies along the wire, points in the direction of the induced electromotive force.

If the student will apply this rule to the case illustrated in Fig. 297, he will find that the induced current is flowing toward him.

The foregoing rule is soon learned by a little practice. Apply it to this case: Consider an east-and-west wire falling to the ground, thus cutting across the earth's magnetic field, which is directed northward. The student is urged to try to apply the rule before reading farther or looking at the answer which is given in the footnote.*

Case 2: a change in the number of magnetic lines threading through a coil. (1) If the magnetic field through a coil is increased, the induced current produces magnetic lines opposite in direction to the increasing lines; (2) if the magnetic field is decreased, the induced current produces lines of force in the same direction as the original field.

The direction of the induced current can be determined from the direction of its field by the right-hand rule (sect. 405).

The rules of Case 2 can be illustrated by simple examples. If the north-seeking pole of a magnet is moved toward a coil, the lines of the magnet will extend through the coil as shown in Fig. 298. The induced current will flow in a direction such that lines will be set up in the coil opposite in direction to those of the magnet. To produce lines in that direction the current, according to the right-hand rule, must be in the direction shown in the figure. If the magnet is now taken away, the induced current will flow in the opposite direction, or in a direction to produce lines through the coil in the same direction as those due to the magnet.

---

*The induced electromotive force will be toward the east.

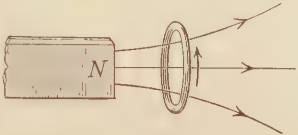


FIG. 298

If a current is started in a coil, there will be an induced current in another coil placed near and parallel to it. In accordance with the rules given above, the momentary induced current in the second coil will flow in a direction opposite to that in the first coil. If the current in the first coil is broken, the induced current in the second coil will be in the same direction as the current in the first coil. The student should verify these statements by applying the rules.

The rules for Case 2 may be summarized by saying that the tendency is for the magnetic field in a coil to stay in the condition in which it is. If the field is increased, lines which oppose the increased field are set up by the induced current; if the field is weakened, the induced current produces lines in such a direction as to strengthen the field.

481. Lenz's law. The following important law was first stated by a Russian, H. Lenz (1804–1865), in 1834: *Whenever an induced current is produced by any motion, the current will flow in a direction such that mechanical forces will be produced which oppose the motion.*

If a magnet is moved toward a short-circuited coil of wire, there will be induced in the coil a current in a direction such that the magnet will be repelled; that is, the magnetic field due to the induced current will repel the magnet (see Fig. 298). On the other hand, if the induced current is produced by taking the magnet away, the current will flow in a direction such that the magnet will be attracted. Thus in either case a force is produced which opposes the motion.

The following experiment illustrates the application of Lenz's law: A sheet or plate of copper is suspended in such a way that it can swing to and fro between the poles of an electromagnet. If, while the copper plate is swinging between the poles, the coils of the electromagnet are connected to a source of current, the plate will be abruptly stopped. If the student tries with his hand to move the copper rapidly back and forth in the strong magnetic field, he will find that a considerable force is required. Whichever way the copper plate is moved, there will be currents induced in it in such a direction as to produce current-magnetic-field forces which oppose the motion.

A silver coin or piece of copper spinning on the end of a thread will be quickly stopped when moved into a strong magnetic field. The induced currents in the coin are always in such a direction that there will be current-magnetic-field forces which oppose the motion of the coin.

The student will now understand why the moving parts of a motor or a dynamo which are in strong magnetic fields should never contain solid masses of metal. To avoid the induced currents and the retarding forces, the revolving iron cores are built up of sheet iron. The film of oxide on the surface of the metal offers sufficient resistance, so that the induced current flowing from one sheet to the next is very small.

Lenz's law can be used in many cases to predict the direction of the induced current. For example, when the north-seeking pole of a magnet is moved toward a coil of wire, the induced current in the coil must flow in such a direction as to produce a magnetic field which opposes the motion of the pole; that is, repels it. It should not be difficult for the student to state, in the case of Fig. 298, which way the induced current must flow in the coil in order to repel the magnet.

Lenz's law is really a corollary to the principle of the conservation of energy. The law can be deduced as follows:

An induced current in any conductor always produces at least a little heat, and hence there is always some expenditure of energy. To supply this energy the external agent which produces the motion, and thus the induced current, must do work. But work can never be done in any motion unless opposing forces are encountered; hence there must be forces which oppose any motion by which induced currents are produced. When a magnet is moved up to a coil, the induced current must repel the magnet. If the induced current attracted the magnet, we should then have the curious and anomalous fact that the current could, without outside aid, increase itself by attracting magnets to it. In the case of the sheet of copper swinging between the poles of a magnet, the kinetic energy of the copper sheet must be sacrificed to obtain the energy of electric currents in the sheet (these currents produce heat; and if the sheet is forced back and forth between the poles, it will get hot). We thus see that if we produce induced currents by any motion, we should expect opposing forces.

482. Relationship between the current-magnetic-field force and the direction of the induced electromotive force. If, as indicated in Fig. 299, a wire perpendicular to the plane of paper is moved downward across a magnetic field by some external force, thus inducing a current in the wire, the magnetic field will be distorted as shown. The method described in section 480 tells us that the induced current in the wire flows toward the student. In section 462 it was shown that such a distorted field would tend to push the wire up. In this case there are at least two forces acting on the wire: the force producing the motion, and the force due to the action of the magnetic field on the current-carrying wire. To distinguish the latter force from others that may be acting on the wire, we have called it the current-magnetic-field force. Fig. 299 shows us that when a current is induced by moving a wire downward through a magnetic field, the current-magnetic-field force will be directed upward; that is, it will oppose the motion of the wire. This is precisely what Lenz's law states.

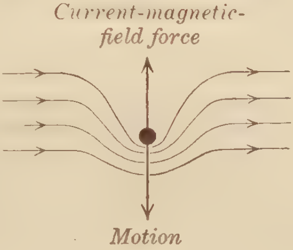


FIG. 299

We thus see that there is a very intimate relation between two phenomena heretofore discussed separately: (1) the direction of the current-magnetic-field force and (2) the direction of the electromotive force induced by the motion of the wire.

Each of the following statements gives a different aspect of the relationship between (1) and (2).

1. If a current is induced by a motion across a field, the induced current flows in such a direction that the current-magnetic-field force will oppose the motion (Lenz's law).

2. If the motion is produced because there is a current flowing in the wire (that is, the motion is caused by a current-magnetic-field force), the induced electromotive force in the wire will be opposite in direction to the current. (The counter electromotive force induced in a motor is an example. See sections 446 and 491.)

The student should carefully verify each of these statements, using the rules for determining the direction of a current-magnetic-field force and the direction of an induced electromotive force.

483. The Arago experiment. Some years before Faraday discovered the principles of induced electromotive forces, Arago made a discovery which was not understood at the time. A horizontal disk of copper was mounted so that it could be set in rotation about a vertical axis. If a horizontal bar magnet was suspended over this disk, the bar would be set in rotation in whichever direction the copper was turning. By some unseen force the magnet was dragged around by the moving plate. Something similar to this is utilized in some types of speedometers, where a metal disk is turned by a rotating magnet.

It was not until after the discovery of induced electromotive forces that Arago's experiment was understood. The rotating copper disk moving through the magnetic field of the bar magnet had electric currents induced in it. Fig. 300 shows the character and direction of these induced currents when the disk is rotating in the direction shown and the bar magnet is above the disk (that is, nearer the student). These currents flow in such a direction that the magnetic field due to the magnet exerts current-magnetic-field forces which tend to stop the disk. The student can verify this from the method of section 462, but it follows at once from Lenz's law. Not only are there forces acting on the disk, but the induced currents produce forces on the poles of the magnet. The magnetic field produced by these currents exerts a torque on the magnet which tends to turn it in the same direction as the disk is moving. (The other torque, which acts on the disk, is in the opposite direction. There is a principle of mechanics, analogous to Newton's third law of motion, which states that for every torque there is an equal but opposite one.)

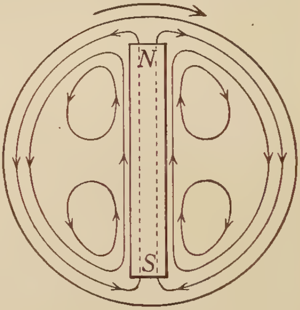


FIG. 300

These electric "whirls" in the copper disk are often called *eddy* currents. The copper plate of section 481, when swinging between the two poles of a magnet, has eddy currents induced in it.

484. The induction coil. Induction coils are used in a great variety of ways: in electrotherapeutics, to explode the gas in gas engines, to operate X-ray tubes, and so on. Fig. 301 shows the electric connections of a simple type. Two coils are wound around the iron core, the primary having only a few turns of wire, and the secondary a great many turns, well insulated from the primary. When a current from the battery, *B*, flows through the primary, the iron core becomes magnetized and attracts the piece of soft iron, *m*. As *m* moves toward the magnet, the contact at *A* is opened, thus breaking the circuit. When the current is interrupted, the iron core loses its magnetism, and *m* is pulled back by the spring on which it is mounted, thus closing the contact so that the current again flows. Just as with an electric doorbell, *m* will continue to vibrate rapidly to and fro, making and breaking the current in the primary. At each make and break of the primary current there is a change in the magnetism of the iron core and an induced electromotive force in the secondary current. The magnitude of the induced electromotive force in the secondary will depend on the *number of turns of wire* in the secondary and on the *rate of change of the magnetic lines* inside it. Many different forms of current-interrupters have been devised which will break the current very quickly. In the form of break shown in Fig. 301, there is a tendency for a spark to form at the place where the circuit is being broken. This spark tends to prolong the life of the current, thus preventing an abrupt break. To remedy this a *condenser* is usually connected in parallel with the break, as shown at *C*. A condenser of the proper size connected in this way

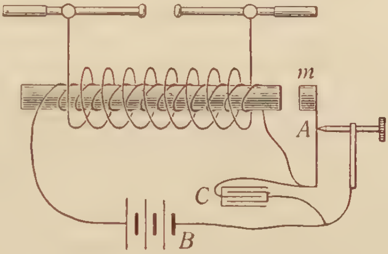


FIG. 301. The induction coil

helps to make a more sudden break of the current, and hence to produce a greater rate of change of magnetic lines. With practically all types of current-interrupters the current changes more rapidly on the break than on the make, with the result that the induced electromotive force in the secondary is always larger on the break of the current than on the make. This is shown in Fig. 302, where the solid lines show the changes in the current in the primary, and the broken lines the induced electromotive force in the secondary. Horizontal distances represent time.

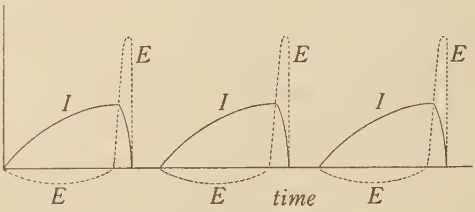


FIG. 302

485. The telephone transmitter. That part of a telephone into which one speaks is called a *transmitter*. The transmitter contains a thin diaphragm, D (Fig. 303), which is set into vibration by sound-waves, and (back of the diaphragm) a large number of small carbon granules, C, loosely packed between two carbon plates. The operation of the transmitter depends on the fact that when the pressure on this granular carbon is changed, the electrical resistance of the carbon changes. The resistance decreases as the pressure is increased, and increases as the pressure becomes less. When the diaphragm vibrates in unison with the sound-waves, it changes the pressure on the carbon. This causes the resistance of the carbon to vary, thus causing the current from the battery to increase and to decrease in unison with the displacement of the vibrating diaphragm. The important purpose of the transmitter is to make an electric current fluctuate in unison with some source of sound.

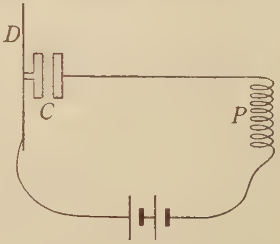


FIG. 303

486. A simple telephone circuit. There are a large number of different kinds of circuits used in telephony. A simple type is shown in Fig. 304. A transmitter, T, and a battery are connected to the primary, P, of a small transformer. The two secondary coils, S (each with a large number of turns of wire), and the receivers, R (see section 405), are connected in series by the transmission lines. The transmitter causes the current flowing through the primary of the induction coil to vary in unison with the sound-waves. The variations in the primary current cause induced electromotive forces in the secondary, and hence a variable current through the receivers. It should not be difficult to see that the diaphragm of a receiver will thus vibrate in unison with the diaphragm of the transmitter, and hence in unison with the sound-waves.

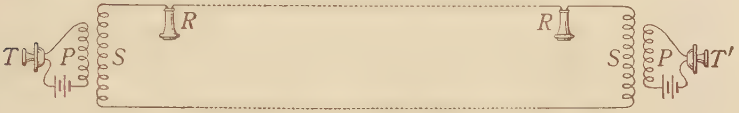


FIG. 304. A simple telephone circuit

In actual practice, telephone circuits are not so simple as indicated in Fig. 304; for example, it is necessary to provide for a method of calling, or ringing a bell. Instead of using a battery at each station, it is now common practice to use one battery at "central" for all lines. This introduces further modifications. But with the exception of very simple cases, induced currents play an important part in all systems.

487. Self-induction. We have seen that there are induced currents in neighboring circuits when the current flowing through a coil increases or decreases. But a changing current induces electromotive forces not only in neighboring circuits but *also in its own circuit*. The reason for this is simple. In the first place, it is necessary to recall that whenever the number of magnetic lines interlinking a circuit changes, there will be an induced electromotive force in that circuit. It makes no difference what the source of these lines may be: they may be those attached to some magnet or they may be produced by some current. In the second place, a current flowing in a circuit produces magnetic lines which interlink its own circuit; when this current increases or decreases, the number of interlinking magnetic lines increases or decreases. Hence an induced electromotive force should be expected.

In the special case where an electromotive force is induced in a circuit by the change of the current in that circuit, it is called an electromotive force of *self-induction*. The reason for this name should become clearer to the student before he has finished reading this section. But he should keep clearly in mind that in this case there need be no primary and secondary coils, for we are concerned now with the electromotive force induced in a coil by changes in the current which is flowing in the coil.

If the student applies a rule for finding the direction of an induced electromotive force, he will see that the electromotive force of self-induction is always in a direction *opposite* to that of the current when the current is *increasing*, and in the *same* direction as the current when the current is *decreasing*. Hence, if an attempt is made to increase or to decrease the current, an *electromotive force is produced which opposes the change*. This is analogous to the inertia of matter: *mass* resists any change in its motion, any attempt either to increase or decrease its velocity. This is the reason for saying that *a current has inertia*. A current in a circuit containing a coil of wire with an iron core has more inertia than one without the iron. Why?

The well-known phenomenon of the *spark at the break* of a circuit is due to self-induction. For example, if a current flowing through a coil of a large number of turns of wire is quickly broken, there will appear a spark at the break. Since the current at the time of the break is changing very rapidly, the magnetic lines which interlink the coil will also change very rapidly. Hence an unusually large electromotive force will be developed in the coil in the same direction as the current—an electromotive force usually large enough to cause a spark to form at the break. When the coil contains an iron core, the energy of the spark at the break is much greater. Why? The spark produced at the break of a circuit that does not contain coils of wire is very small. On short-circuiting a dry cell, and then breaking the circuit, a much smaller spark is obtained than when a small electromagnet is included in the circuit.

The connections for an experiment commonly shown in classroom demonstrations to illustrate self-induction are given in Fig. 305. A direct current is sent through the coils, *M*, magnetizing the iron core. When the key, *K*, is opened, and the current suddenly decreased, the self-induced electromotive force in *M* will cause the lamp, *L*, to flash up brightly. The lamp may not flash on the make of the current, although it flashes brightly on the break. Why?

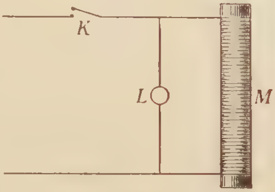


FIG. 305

If a bank of lamps, connected in parallel with each other, and a coil of wire with a removable iron core are connected in series to an alternating-current source, the lamps may glow brightly when the iron core is out of the coil. But inserting the iron core dims the lights, owing to the increased self-induction of the coils. Since the alternating current is continually changing, the magnetic field produced in the coils by this current is continually changing, thereby inducing electromotive forces which oppose the change. Since the current in an alternating-current circuit is continually changing, the electromotive force of self-induction is always present and plays an important part.

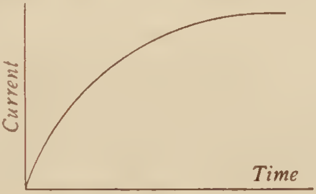


FIG. 306

On account of self-induction the magnitude of an alternating current is usually less than that predicted by Ohm's law. Hence special methods must be used in computing the current in an alternating-current circuit. However, in those cases where the current does not flow through coils, and hence does not produce strong magnetic fields, the effect of self-induction is small.

In Fig. 306 is shown the retarding effect of self-induction in the starting of a direct current through an electromagnet. In the case of large electromagnets having very many turns of wire it may be a matter of several minutes before the current reaches its maximum value.

488. Inductance, or coefficient of self-induction (L). The magnitude of the induced electromotive force of self-induction in any circuit depends not only on the rate of change of the current but also on the nature of the apparatus in the circuit. For example, if the circuit contains a coil of many turns of wire, the self-induction will be greater than if the coil contains only a few turns. Or, as has been shown, if the coils have iron cores, larger effects are produced than if there is no iron. For purposes of comparing the effects of different kinds of apparatus the term *inductance* or *coefficient of self-induction* is used.

In order to understand the meaning of this coefficient the expression "interlinkage of magnetic lines through a coil" must be understood. Interlinkage is the total number of magnetic lines threading through any coil multiplied by the number of turns of wire in the coil; in other words, if there are N turns of wire, each interlinking line is counted N times. For example, if there are 100 magnetic lines going through a coil of 25 turns, the interlinkage is 2500.

If the coil does not contain an iron core, the number of magnetic lines interlinking it is proportional to the current which produces the lines. This may be expressed by the equation

$$\text{Interlinkage} = LI, \tag{81}$$

where L is a constant for any given coil. This constant is called the *coefficient of self-induction* or, more simply, the *inductance*. From equation (81) the coefficient, L, may be defined as equal to the *interlinkage produced by unit current flowing through the coil.*

An important and commonly used expression involving the use of inductance can be deduced as follows:

The induced electromotive force in any coil is always equal to the rate of change of the interlinking lines through that coil; or, as the interlinkage is equal to LI,

$$\text{E.M.F. of self-induction} = \text{rate of change of } LI.$$

Since L is a constant,

$$\text{E.M.F. of self-induction} = L \times \text{rate of change of } I. \tag{82}$$

If the electromotive force is measured in volts, and the current in amperes, the inductance in equation (82) is measured in *henries*.* For convenience the millihenry is commonly used.

$$1 \text{ henry} = 1000 \text{ millihenries.}$$

---

*Named after Joseph Henry.

489. Energy in a magnetic field. There are excellent reasons for believing that there is energy stored in a magnetic field; that is, energy distributed throughout the entire field. Where the field is the strongest—that is, where there is the greatest density of the magnetic lines—there is the most energy.

This notion can be used to explain many things about electric circuits that produce strong magnetic fields. For example, we have seen that when a circuit is closed, the current requires some time to rise to its full value (sect. 487). Knowing that this current produces a magnetic field which has energy, we should expect this. For a finite amount of energy cannot be given a field instantaneously. Time is required, and we should expect that some time would elapse before the current and its magnetic field reach the final value. There is another point of view. When the current is increasing, there is an opposing electromotive force. In section 446 it was shown that work is done whenever there is an opposing electromotive force. In this case what is the result of the work? It is this work that supplies the energy in the magnetic field.

The spark at the break of a circuit which has an electromagnet in it contains an appreciable amount of energy. This energy comes from that stored in the magnetic field. In the experiment of section 487, where a lamp flashes when the circuit is broken, the energy in the flash comes from the energy of the magnetic field.

In the transformer the energy is in some way carried from one circuit to another. It may be said that the primary current supplies energy to the magnetic field, which in turn gives it to the secondary.

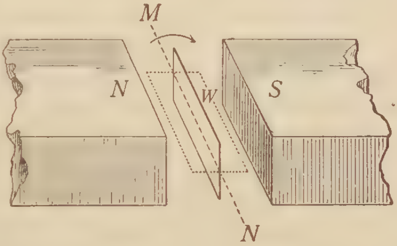


FIG. 307

490. Dynamos. The most important type of electric generator, the dynamo, depends on the induced electromotive force that is produced when conductors move across magnetic fields. We shall consider first the case of a rectangular coil of wire that can rotate on the axis MN (Fig. 307). As the coil rotates, in the indicated direction, the top wire (marked W in the figure) will move down across the field, inducing an electromotive force directed toward the student. The induced electromotive force will continue in this direction until the wire W reaches the lowest point (C, Fig. 308, where a vertical section of the coil is shown). The largest electromotive force will occur while the wire is cutting the lines at the greatest rate; that is, at the point *B* (Fig. 308). From *C* around through *D* the electromotive force in the wire *W* will be reversed. Fig. 309 shows the relative direction and magnitude of the induced electromotive force when the wire *W* is moving through the positions *A*, *B*, *C*, and *D* of Fig. 308. Distances above the horizontal line refer to an electromotive force in one direction, while those below refer to a reversed electromotive force. When the wire *W* is at *A*, it is moving parallel to the field, and the rate of cutting lines is zero; hence the electromotive force is zero. At *B* the rate is the maximum; for the wire is there moving at right angles to the field, and cutting the lines at the greatest rate. At *C* the electromotive force reverses in direction, because the wire *W* begins to move in the opposite direction across the field. The student should follow this coil through a complete revolution and verify the graph of Fig. 309.

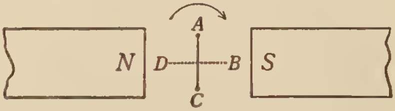


FIG. 308

If a coil of this kind is connected to an external circuit by a commutator (see Fig. 289, p. 490), the current in the external circuit will always flow in the same direction.

A single coil like this would give a small electromotive force; but, by winding on many turns, the electromotive force can be made fairly large. Usually a large number of coils is wound on an iron core, just as in the case of a direct-current motor. In fact, any direct-current motor can be used as a dynamo.

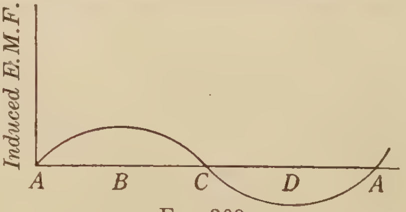


FIG. 309

The magnetic field in which the coil or armature rotates is produced by an electromagnet, the current being obtained from the rotating armature. It is easy to understand how the magnetizing current for the field magnets can be supplied by current from the armature after the machine is once "excited." But in order to have induced currents, one must first have a magnetic field. How can the machine "pick up," or begin generating current, when it is started? The iron cores of the large electromagnets used in dynamos always contain enough "residual" magnetism to give a small induced electromotive force when the armature is rotated. When the small current produced by this electromotive force flows through the wire wound around the electromagnets, it will strengthen the magnetic field; and this, in turn, will produce a larger induced electromotive force in the rotating armature, and this a still larger current in the field coils, and so on. Thus the electromotive force rises quickly to its final value.

The field windings may be connected to the armature in several different ways; for example, they may be in *series* or in *shunt* with the armature. A common type is a combination of series and shunt windings, called the *compound* winding. These methods were illustrated in section 470, in the discussion of the motor.

491. The counter electromotive force of a motor. It was shown in section 446 that a motor must have a counter electromotive force if energy other than that used in heat is given to it. At that time we might have anticipated some of the results of this chapter. For we might have guessed that the motion of the coils of wire across the magnetic field was what produced the electromotive force, and from this we might have inferred that an electromotive force would be induced at any time that wires cut across a magnetic field.

492. The alternating-current dynamo. The alternating-current dynamo, or generator, is used commonly for generating electrical energy which must be distributed over a considerable area, such as a city. There are two principal parts to a generator: the armature, which carries the coils in which the electromotive force is induced, and the "fields." It is now common practice for the armature to be stationary and the fields to revolve. Fig. 310 shows diagrammatically a generator of this type. As shown, there are a number of magnetic poles arranged with alternate polarity. These are marked S, N, S, N, and so on. These poles are magnetized by a direct current, usually obtained from a small separate direct-current generator called the *exciter*. The armature coils, in which the electromotive force is to be induced, are wound on stationary soft-iron poles mounted on the frame of the machine just outside the revolving poles, and connected in series.

When an *N* pole stands in front of a particular coil on the stationary armature, the magnetic field from that pole magnetizes the iron around which the coil is wound, thus producing a strong magnetic field through the armature coil; but when an *S* pole stands in front of this coil, the magnetic field through it is in the opposite direction. Thus as *N* moves away and *S* approaches, an induced electromotive force is produced in the coil. When *S* is leaving, and another *N* is approaching, the induced electromotive force in the coil is in the opposite direction. Hence, as the poles revolve past this coil, an alternating electromotive force is produced in it. The same thing is true of each one of the armature coils.

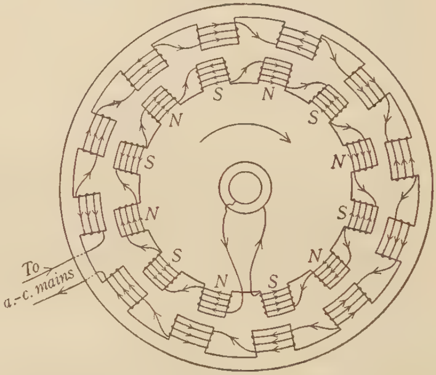


Fig. 310

It should be readily seen that the electromotive forces induced in adjacent armature coils at any instant will be in opposite senses; but, by winding the adjacent coils in opposite directions and then connecting them in series, the electromotive forces add, and produce for the entire armature an electromotive force which may be relatively large (the direction of the induced electromotive force is indicated in the figure on each of the armature poles). Fig. 311 shows a

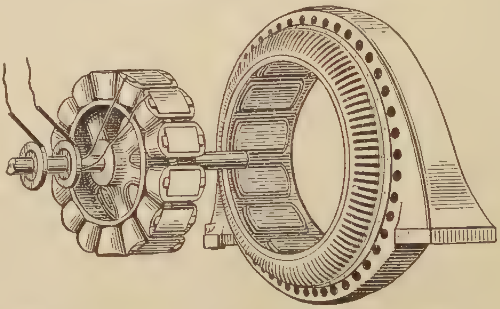


Fig. 311

revolving-pole machine, with the revolving poles partially removed so as to expose part of the armature windings.

What has Lenz's law to do with the fact that an engine or a turbine is needed to drive a dynamo?

493. Rotating magnetic fields: the induction motor. The most common type of alternating-current motor, the *induction* motor, depends for its action on the fact that a rotating magnetic field can be produced by alternating currents. To understand this it is necessary to know what is meant by the term *phase* as applied to oscillating motions. As a simple example take the case of two pendulums. They may both have the same period, or time of swing, and they may both have the same amplitude; yet they may not have the same motion, for one might be moving to the left while the other was swinging to the right. *Phase* is a term that states in what part of its oscillation the pendulum may be and in what direction it is moving at any instant. If the two pendulums are not swinging together, their phases are not alike. Take the case of two alternating currents, shown by the curves of Fig. 312. Vertical distances indicate the magnitude of the current; horizontal distances, time. When the curve is above the straight horizontal line, it indicates that the current is flowing in one direction; when the curve is below, the current is flowing in the opposite direction. If the two curves A and B represent two different alternating currents, it is clear from the figure that they do not reach their maximum values at the same time, nor do they reverse at the same time. Their *phases* are different. A combination of circuits carrying two currents like these is usually called a *two-phase* circuit.

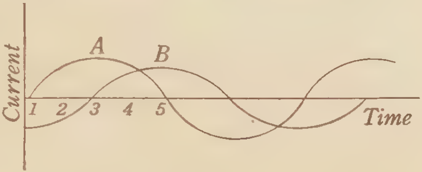


FIG. 312

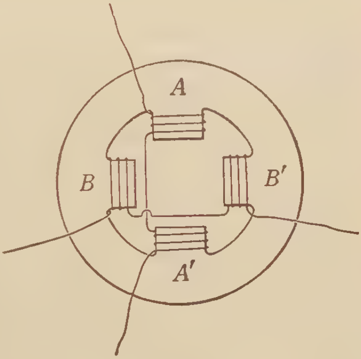


FIG. 313

We shall now see how it is possible to get a *rotating* magnetic field with a two-phase supply. Fig. 313 shows an iron frame on the inside of which are mounted four poles. The windings of two of these poles, A and A', are connected to one of the alternating currents, while those of the poles B

and B' are connected to the other current. Since these two currents are not in phase, the two pairs of poles will not reach their maximum strength at the same time. Fig. 314 shows the polarity and the direction of the magnetic field (flux) at five different times corresponding to the numbers in Fig. 312. During this interval the resultant magnetic field between the poles has changed in direction and has rotated counterclockwise through 180 degrees. In one complete cycle it will make a complete rotation, 360 degrees. Hence if two alternating currents, differing in phase, flow around the poles of an electromagnet similar to that of Fig. 313, a rotating magnetic field will be produced.

If any solid conductor, such as a sheet of copper or aluminum, is placed in a rotating field, there will be induced currents produced in it by the moving magnetic field. There will thus be a current flowing in a conductor which is in a magnetic field. This gives the necessary conditions for the existence of a current-magnetic-field force. This force will be in such a direction as to tend to drag the metal around with the moving field. Whenever currents are induced, there are forces which oppose the *relative* motion (Lenz's law). When an armature which has all its coils short-circuited is placed in a rotating field, the induced currents will be large enough to produce large torques. Thus a rotating magnetic field causes the armature to rotate.

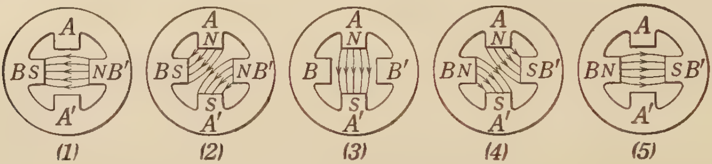


FIG. 314

Since the armature of such a motor has only induced currents in it, this type is called an *induction* motor. The armatures are usually very simple in construction, consisting of short-circuited bars of copper placed in slots in a cylindrical iron core. No commutator or brushes are necessary, for the armature does not need to be connected to an external circuit.*

The standard practice in this country is to transmit electric power by alternating currents over three wires. Three wires are used in order to get currents having different phases, so that induction motors can be used.

494. Long-distance transmission of energy. We can now understand many of the details of long-distance transmission of electrical energy. Since supplies of energy (from waterfalls or coal mines, for example) are usually not located where the energy is needed, some method must be used to transport that energy. Electrical methods not only have proved to be the best but often are the only feasible ones.

---

*If possible the student should see an induction motor which has been partially dismounted.

A brief description of the more important parts of an electrical installation will be given. If water power is available, the water is used to run turbines, which in turn operate alternating-current dynamos, where the mechanical power is converted into electrical power. If some form of fuel is to be used, steam must be generated, and engines or steam turbines must be driven with the steam. These in turn drive the dynamos. From the dynamos the power is delivered to step-up transformers, which may raise the voltage to 50,000, 60,000, or even 200,000 volts. At high voltages the power is transmitted with relatively small heat losses in the transmission circuit; for when the voltage is high, the current may be small, and hence the I^2r loss be small. At the other end of the line the power is delivered to step-down transformers, the low-voltage side giving any voltage that may be wanted. It is not necessary that all the power be delivered to one transformer; a number may be connected, in parallel, across the high-voltage line. Hence part of the power may be converted into a low-voltage, large-amperage circuit at any one place, while the rest may be transmitted farther at the high voltage. It may not be advisable to transform the energy in one step to a low potential. For example, it would not be feasible to transmit very high voltages through a city. Hence in the suburbs the voltage may be reduced by a transformer to 2200 volts, the power being transmitted at this voltage to places where the voltage may be reduced still farther by other transformers. The high efficiency of a transformer makes the loss due to the use of a number of them small. It is apparent that power transmitted by alternating currents becomes, through the aid of transformers, very flexible, for the voltage and amperage can be varied at will.

Power is now transmitted over rather long distances; for example, a large amount of power from the Mississippi River at Keokuk, Iowa, is transmitted to St. Louis, a distance of one hundred and forty-four miles. There are now many instances of long-distance transmission of power by electrical means.

495. Summary. An induced electromotive force is produced in a coil while the magnetic field through the coil is increasing or decreasing. Increasing the magnetic field produces an electromotive force in one direction, and decreasing it produces an electromotive force in the opposite direction.

An induced electromotive force in any coil is equal to the rate of change of the number of magnetic lines in the coil multiplied by the number of turns of wire. Its average value is equal to

$$N\frac{\phi_2 - \phi_1}{t}, \tag{80}$$

where ϕ_1 is the number of magnetic lines threading through the coil at the beginning, and ϕ_2 the number at the end of the interval of time, t.

Another and equivalent way is to say that whenever magnetic lines are cut by a conductor, an electromotive force is induced, and that its magnitude is equal to the rate of the cutting of magnetic lines by a conductor.

An application of induced electromotive forces is made in the transformer, which is used to step the voltage of an alternating current up or down, as needed. When the secondary has a larger number of turns than the primary, the induced electromotive forces are larger than the electromotive forces in the primary circuit, and the transformer is called a *step-up* transformer; if the number of turns in the secondary is smaller, it is called a *step-down* transformer.

Lenz's law states that whenever an induced current is produced by any motion, the current will flow in a direction such that mechanical forces will be produced which oppose the motion.

The student should become familiar with methods of determining the direction of an induced electromotive force. Whenever it is easy to see that the induced electromotive force is due to a conductor moving across magnetic lines, or to lines moving across a conductor, the distorted-field method of section 480 is probably the most convenient; but if the case is one of a change in the number of magnetic lines in a coil, Case 2, sect. 480, is probably the best method. Lenz's law frequently gives a convenient method of determining the direction.

In the disk of Arago, in the induction coil, and in telephone circuits there are applications of induced electromotive forces.

Self-induction is an important case of an induced electromotive force. Whenever the current in a circuit is increased or decreased, there appears an electromotive force of self-induction that opposes the change. A large number of phenomena are due to self-induction. In the case of an alternating current—where the current is continually changing—self-induction plays a very important part. Usually the current in such a circuit cannot be computed by Ohm's law when the voltage and resistance are known.

The interlinkage of a coil is equal to the total number of lines threading through the coil multiplied by the number of turns of wire. If the coil does not contain an iron core,

$$\text{Interlinkage} = LI, \tag{81}$$

where L is a constant called the coefficient of self-induction. In any circuit

$$\text{E.M.F. of self-induction} = L \times \text{rate of change of } I. \tag{82}$$

If the electromotive force is in volts and the current in amperes, L is measured in *henries*.

The common method of developing electrical energy is by the use of either the direct-current dynamo or the alternating-current dynamo. Both these types utilize induced electromotive forces.

By using alternating currents of more than one phase, rotating magnetic fields can be produced. Any conductor placed in a rotating magnetic field has induced currents in it. There is thus produced a current-magnetic-field force that produces a torque tending to rotate the conductor in the direction in which the magnetic field is rotating. The most common type of alternating-current motor is the induction motor. The armature of this type is placed in a rotating magnetic field, which induces currents in it. Since these currents are in a magnetic field, there are current-magnetic-field forces acting which produce rotation of the armature.

PROBLEMS

1. The north-seeking pole of a bar magnet is plunged from above into a coil lying on a table. Is the induced E. M. F. in the coil, as one looks down, clockwise or counterclockwise?

2. The plane of a coil of 200 turns, and 200 sq. cm. in area, is perpendicular to a magnetic field the strength of which is 100 C.G.S. units. Compute the E.M.F. in volts when this field is cut out in one tenth of a second.

3. 50,000 magnetic lines of force are cut out of a coil of 100 turns in 1/100 sec. Compute the mean E.M.F. in volts.

4. A circular coil, of radius 20 cm. and with 200 turns, lying on the floor, is turned over in 0.2 sec. Compute the average E.M.F. in volts if the vertical component of the earth's field is 0.55 C.G.S. unit.

5. A coil (Fig. 315) is rotated continuously about an east-and-west axis. (Assume the earth's field to be horizontal.) (*a*) In what direction is the induced E.M.F. on the side of the coil that moves downward? (*b*) At what places in its rotation is the E.M.F. a maximum? (*c*) a minimum?

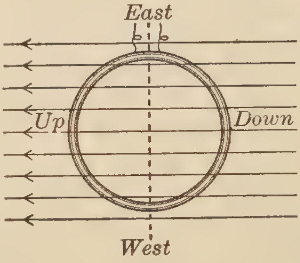


FIG. 315

6. An east-and-west wire falls to the ground with an average speed of 250 cm./sec. If the horizontal component of the earth's magnetic field is 0.20, find the E.M.F., in volts, induced in 100 m. of the wire.

7. A wire 10 cm. long moves across a uniform magnetic field of 10,000 lines per square centimeter at such a rate that an E.M.F. of 0.5 volt is induced in the wire. Compute the speed of the wire.

8. What is the induced E.M.F. in volts in the iron axle of a train traveling northward at 60 mi./hr.? (The length of the axle is 4 ft. 8 in., and the vertical component of the earth's magnetic field is 0.55 C.G.S. unit.)

9. The primary of an induction coil produces 50,000 magnetic lines. If this field disappears in 0.01 sec., how many turns must the secondary winding have to produce 10,000 volts?

10. The secondary of an induction coil has 200,000 turns. In how short a time must 100,000 magnetic lines be removed from this coil to produce 15,000 volts?

11. The north-seeking pole of a magnet is held above an east-and-west wire. (*a*) If this wire is moved sidewise northward, what will be the direction of the induced E.M.F.? (*b*) If the wire had a current flowing eastward, what would be the direction of the current-magnetic-field force acting on the wire? (Neglect the effect of the earth's field.)

12. A coil of 200 turns and an area of 200 sq. cm. is revolving at the rate of 50 revolutions per second about an axis perpendicular to a magnetic field the strength of which is 4000 C.G.S. units. What is the average E.M.F., in volts, when the coil rotates through 180°, starting at a point where the coil has a maximum number of lines passing through it?

13. A metallic door in an east-and-west wall is hinged at the west edge. It is opened southward and swung through 180°. (*a*) What is the direction of the induced currents in the door? (*b*) At what point of the motion will the induced E.M.F. be the largest?

14. How many revolutions per minute must a 40-pole alternating-current generator have, in order to produce a 60-cycle current?

CHAPTER XXXIV

MAGNETIZATION OF IRON

Introduction, 496. Magnetizing force (H), 497. Magnetic induction (B), 498. Curves of magnetization, 499. Magnetic permeability, 500. Diamagnetic substances, 501. Magnetic hysteresis, 502. Magnetic circuits, 503.

496. Introduction. On account of the large number of practical applications of magnetism of iron, a knowledge of the simpler facts of the behavior of iron in magnetic fields is important. Certain concepts are used in practically all the literature pertaining to the magnetic properties of iron and of other magnetic substances. Such terms as *magnetic induction*, *magnetizing force*, and *permeability* are commonly employed. The meaning of these terms can be more clearly understood by describing some experiments with induced electromotive forces.

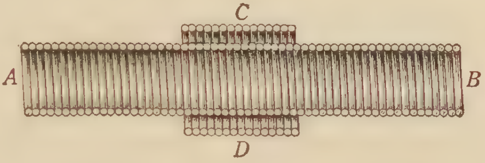


FIG. 316

Let us consider a long solenoid, or helix, AB of Fig. 316, around which is another winding, a secondary, CD. When a current flows through the primary, AB, magnetic lines are produced which run through the solenoid and, at the same time, interlink the secondary coil, CD. Whenever the current in AB is made or broken, the changing magnetic field produces in CD an induced electromotive force, and hence a momentary current in the circuit (not shown in the figure) connected to CD.

It can be proved* that the *total* quantity of electricity, Q (average current times the time of flow), flowing in the secondary circuit when the primary current is made or broken is given by the expression

---

* A simple proof is given here. Φ is the flux, or the total number of magnetic lines cutting across the secondary coil when the current in the primary is made or

$$Q = N_2 \frac{\Phi}{R_2}, \tag{83}$$

where R_2 is the resistance of the secondary circuit, N_2 the number of turns of wire in the secondary, and Φ the *flux*, or the total *number of magnetic lines* running through the primary when the current is flowing in it.

It is not difficult to measure by a *ballistic* galvanometer[*] the quantity, Q, flowing in the secondary. The number of turns, N_2, and the resistance, R_2, of the secondary circuit can also be found. Hence everything in equation (83) can be measured except Φ. Since it is the only unknown quantity, its value can be easily computed. By a simple transformation of this equation the value of Φ is given by

$$\Phi = Q \frac{R_2}{N_2}. \tag{84}$$

The reason that equations (83) and (84) have been stated here is to point out that we have available an *experimental method for determining flux, or the number of magnetic lines.*

Let us suppose that the flux through the helix has been measured in this way. If iron is placed inside the helix, and the experiment

---

broken. When the primary current is broken, there is induced in the secondary an average electromotive force equal to $N_2 \frac{\Phi}{t}$, where N_2 is the *number of turns* in the secondary coil, and t is the length of time the lines are cutting across the secondary.

$$\text{Average current in the secondary} = \frac{\text{average E.M.F.}}{R_2}.$$

Substituting in this the value of the average electromotive force,

$$\text{Average current} = N_2 \frac{\Phi}{R_2 t}.$$

But $$Q = \text{average current} \times t.$$

Hence $$Q = N_2 \frac{\Phi}{R_2}. \tag{83}$$

Both Q and R_2 must be measured in C.G.S. electromagnetic units.

[*] A ballistic galvanometer is one where the first "throw" is used — the maximum, not the steady deflection.

repeated, a much larger induced electromotive force is produced in CD, and a larger Q flows. This shows that the magnetic flux is much greater when iron is present. By the aid of equation (84) we can then compute the flux in the iron, or, in other words, determine the magnetic condition of the iron. This is perhaps the simplest method of doing it. In the following sections this experimental method will be made the basis for explaining the meaning of the magnetic terms commonly used.

497. Magnetizing force (H). The strength of the magnetic field at any place is equal to the force acting on a unit north-seeking magnetic pole at the place in question. This definition is, by its nature, limited to a field in air. For example, in the solenoid referred to in the last section, this definition can be directly applied to a point in the field inside the solenoid *when there is no iron* at that point. It has been pointed out (sect. 365) that in air the number of magnetic lines per square centimeter is always numerically equal to the strength of the field. For example, if one should take the total number of lines running through the primary (as found by the experiment of the last section) and divide it by the sectional area of the core of the primary in square centimeters, he would get the number of lines per square centimeter. *If there were no iron in the core*, this number would be equal to the strength of the magnetic field in the solenoid. The strength of the field is often called the *magnetizing force*, and is denoted by the letter H.

Let us now consider the case where iron is placed inside the secondary. The magnetic flux, or the total number of magnetic lines, would be greatly increased; but what we have called the *strength* or *magnetizing force* of the field would be the same as before, for it depends on the current, which has not been changed. Hence we may say that *the magnetizing force, H, in the iron is equal to the number of magnetic lines per square centimeter which would be present if the iron were absent.**

---

*This is not accurate for cases where a short piece of iron is used. The ends of all pieces of iron have poles when magnetized, and these poles usually tend to decrease H in the iron. The poles of short pieces of iron have a *demagnetizing* effect. In cases where the ends are far apart or where there are no ends (and hence no poles, as in transformers) the statement is exact.

498. Magnetic induction (B). When the iron is placed in the solenoid, the iron becomes magnetized, adding to the lines produced by the current the lines from its own magnetism. The number of lines per square centimeter is no longer numerically equal to the strength of the field, or the magnetizing force H, but to a new quantity called the *magnetic induction*—a quantity which may be represented by the letter B.

The total number of lines, Φ, running through the iron can be determined by the experiment of section 496 and by equation (84). Since B is the number per square centimeter, we have

$$B = \frac{\Phi}{\text{cross section of iron core}}. \tag{85}$$

The magnetic induction B is the flux per square centimeter*, the area being perpendicular to the flux.

In a transformer the magnetizing current in the primary produces a certain magnetizing force, H; the iron becomes magnetized, giving a certain magnetic induction, B. It follows from the definitions given that the induced electromotive force in the secondary depends directly on the time rate of change of B and not on that of H. The same is true of all other cases where iron is used in producing induced electromotive forces.

All magnetic lines are continuous; that is, they have no ends. In an iron rod or bar magnet they run out into the air near one end and return through the air to the other end and then back through the iron. In air there is really no numerical difference between H and B, both being equal to the number of magnetic lines per square centimeter. But inside of iron, H is

---

*No precise definition of magnetic flux, Φ, has been given in this chapter. It is, as appears from section 496, measured in terms of the induced electromotive force it will produce. The first paragraph of the footnote of that section contains a definition of Φ. From this paragraph we obtain the equation

$$\frac{\Phi}{t} = \frac{\text{induced E.M.F.}}{\text{number of turns}}.$$

Or, the magnetic flux is that quantity the rate of change of which in a coil of one turn is equal to the induced electromotive force in that coil. The flux Φ is equal to what has been called in other chapters the *number of magnetic lines*.

different from B and therefore is not equal to the number of lines per square centimeter.

499. Curves of magnetization. By means of the experiment given in section 496 it is not difficult to find the relation between the magnetizing force, H, and the magnetic induction, B, of a given sample of iron. The student should remember that H is the number of magnetic lines per square centimeter in the core when the iron is not present, and B the number when the iron, in the form of a long bar, is inserted in the solenoid.

The value of H is proportional to the current in the primary; and it is possible and also quite common to compute H from the value of this current and the dimensions of the primary. Since H is proportional to the current, any desired value of H can be obtained by changing the current.

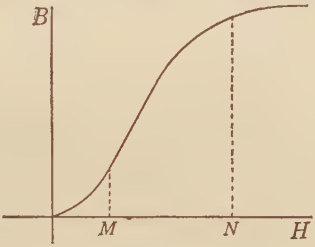


FIG. 317. A magnetization curve

Fig. 317 shows the values of B, where the specimen has been subjected to increasing values of H. A curve of this kind is called the *curve of magnetization* of the particular kind of iron used (the figure is somewhat misleading, for the values of B and H are not plotted to the same scale, the values of B being much larger than those of H). It will be noted from the curve that for very small values of H an increase in H has only a small effect on the value of B; but after H has reached a value corresponding approximately to M in the figure, the magnetic induction increases very rapidly with an increase in H. For much larger values of H the curve ceases to rise so rapidly; and after values of H larger than N are reached, there is relatively little increase in B, even for very large changes in H. The iron is then said to be *saturated.*

The curves for different types of iron and steel are different, but all have the same general form as the curve in Fig. 317. The curve for wrought iron rises much more rapidly and reaches saturation for much smaller values of H than the one for cast iron. Different kinds of steel show similar differences.

500. Magnetic permeability. *Magnetic permeability* (μ) is defined as the ratio of the magnetic induction B to the magnetizing force H;

that is,
$$\mu = \frac{B}{H}, \tag{86}$$
or
$$B = \mu H. \tag{87}$$

A knowledge of the permeability, μ, enables one to compute the number of lines produced in iron for a known value of H.

The permeability of iron is not a constant, but is different for different values of H. This can be readily seen from the curve of Fig. 317; for if μ were constant, B would be proportional to H, and the curve would be a straight line.

501. Diamagnetic substances. Some materials (for example, bismuth, antimony, and, to a lesser degree, many common substances) have a permeability which is slightly less than that of air; hence magnetic lines prefer to go through air rather than through these substances. Such substances are called *diamagnetic*, because, when they are placed near the pole of a strong magnet, they tend to turn so that their length is across the magnetic field. In section 358 it was stated that these substances are feebly repelled when brought near the pole of a strong magnet.

502. Magnetic hysteresis. In any alternating-current electromagnet, iron is magnetized first in one direction and then in the other, as the magnetizing current changes from one direction to the other. It is a very interesting and important fact that the magnetization of the iron does not reverse at exactly the same time that the current does, but *lags behind*. This is shown in Fig. 318, where horizontal distances are proportional to the magnetizing current, and vertical distances are proportional to the magnetic induction, B. At the point a the current is at a maximum value in one direction, and B is also at its maximum value. At b the current has been reduced to zero, but B is nearly as large as before. The distance Ob indicates the *residual magnetism* of the iron. It is only after the current is reversed and then increased until it reaches a value Oc that B falls to zero. After that, B increases very rapidly in the opposite direction, reaching the other maximum at the same time that the current does, as is

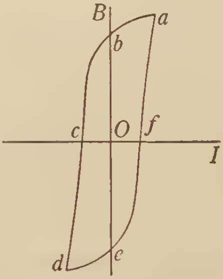


FIG. 318. A hysteresis loop

shown by the point d on the curve. When the current is again reduced to zero, the iron retains, as before, much of its magnetization (the value of B is shown by the distance Oe). At the point f a current Of is flowing, but the iron has entirely lost its magnetism. When the current again reaches its maximum, however, the iron becomes fully magnetized, as is shown by point a on the curve. This lag in magnetic properties behind the varying current is called *hysteresis*.

The lag, or hysteresis, is well shown in Fig. 319. The smooth curve represents an alternating current used to magnetize the iron, while the broken line is the resultant magnetic induction. Horizontal distances are proportional to *time*. The letters on the horizontal axis refer to the same states as are represented by the corresponding letters in Fig. 318.

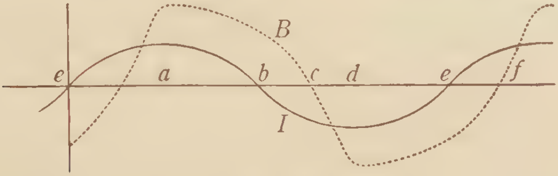


FIG. 319

503. Magnetic circuits. It was stated in section 498 that magnetic lines form closed paths; the lines are endless. Another method of stating this is to say that the lines always form a *closed magnetic circuit*.

In an iron ring magnetized by a current flowing through wires wound on the ring, the lines are in iron throughout their path, and the magnetic circuit is entirely in the iron. Transformers are usually constructed with "closed" magnetic circuits. In such cases there are no magnetic poles in the magnetic circuit.

Advanced students in engineering often find it advantageous to think of this magnetic circuit as if something actually flowed around it. Certain methods and terms used by engineers are based on this notion; for example, the name *flux* originated in this way. It should be clearly understood that these methods are used merely for the sake of clearness and simplicity of thought and not because anyone believes that such a flow actually takes place.

PROBLEMS

1. In a sample of cast iron, when H was 4 C.G.S. units, B was 550 C.G.S. units; and when H was 8 C.G.S. units, B was 2100 C.G.S. units. Compute the permeability in each case.

2. In a sample of wrought iron, when H was 2 C.G.S. units, B was 2200 C.G.S. units; when H was 4 C.G.S. units, B was 6800 C.G.S. units; when H was 6 C.G.S. units, B was 8500 C.G.S. units; and when H was 10 C.G.S. units, B was 10,300 C.G.S. units. Compute the permeability for each case.

3. A long iron core of the primary of an induction coil has a cross-sectional area of 15 sq. cm. When the current produces a magnetizing force of 6 C.G.S. units, what will be the total flux in the iron if the corresponding value of the permeability is 1500?

4. The maximum permeability of a sample of nickel was 200. This was for the case where H was equal to 2 C.G.S. units. Compute the corresponding value of B.

5. The total flux out from a magnet is $4\pi m$ C.G.S. units, where m is the pole strength. (a) Compute the flux if m is 200 C.G.S. units. (b) Find B inside the magnet if the cross section is 2 sq. cm.

CHAPTER XXXV

POTENTIALS OF CHARGED CONDUCTORS; CAPACITANCES OF CONDENSERS

Summary of facts previously explained, 504. Potentials of charged isolated conductors, 505. The effect of another charged body on the potential of a conductor, 506. A grounded conductor brought near a charged conductor, 507. The condenser, 508. A condenser connected to an alternating-current circuit, 509. Quantitative relation between the charge and the voltage of a condenser, 510. Capacity, or capacitance, 511. Capacitance dependent on physical dimensions, 512. Capacitance dependent on the medium between the plates, 513. Units and formulas for capacitance, 514. Condensers in parallel and in series, 515. Energy of a charged condenser, 516. Summary, 517.

504. Summary of facts previously explained. It has been shown that in order to have a current flow from one conductor to another, the two must be at different potentials. But there is another important effect of a difference of potential. When two conductors are at different potentials, they will have *static* charges of different kinds of electricity on their surfaces. Even if the two conductors are connected by a wire, these static charges will exist if there is a difference of potential. There are many cases where these static charges cause important phenomena. Before proceeding to discuss these cases, however, it is important for the student to be certain that he knows the following facts, which have been previously explained:

1. A positive charge tends to move from a region of higher potential to one of lower potential.

2. A negative charge tends to move from a region of lower potential to one of higher potential.

3. If there is a difference of potential between any two points which are connected by a metallic conductor there will be a current flowing from one to the other; if there is no current, there is no difference of potential.

4. We are usually concerned only with the difference of potential between two different conductors and not with their absolute potentials. This difference is commonly measured in volts.

5. The difference in potential between two points is equal to the work done by some external agency in carrying a unit positive charge from the point of lower potential to that of higher potential.

505. Potentials of charged isolated conductors. It is customary to consider the potential of the earth as zero, and to call all higher potentials *positive*, and all lower potentials *negative*. For example, if an *isolated positively* charged conductor should be connected to the earth, positive electricity would tend to flow to the ground, or a negative charge would tend to flow from the earth to the conductor; hence, according to paragraphs 1 and 2 of the last section, such a conductor has a potential higher than that of the earth. Being at a higher potential than the earth, it is said to have a *positive* potential. On the other hand, in the case of an *isolated negatively* charged conductor, negative electricity would tend to flow from the conductor to the earth, if it were connected to the earth. Hence, according to paragraph 2 of the last section, it has a potential lower than that of the earth. Its potential is *negative*. The terms *positive* and *negative* are used here in a manner similar to their use in connection with thermometer readings. All potentials above zero are called positive, while all potentials lower than zero are called negative.

The student should see that the term *potential*, as used here, really refers to a *difference* of potential—the difference between the potential of the conductor and that of the earth.

506. The effect of another charged body on the potential of a conductor. Let B be a conductor with no charge on it; that is, at the same potential as the earth, or *zero* potential. Now bring near B a positively charged conductor, A. A will act by induction on B, producing the distribution of charges shown in Fig. 320 and explained before (sect. 378). What change has taken place in the potential of B? The student can determine this if he can answer the following question: If B is connected to the earth, in what direction will the current flow? If he knows the answer, then, by applying paragraph 1 or 2 of section 504, he can decide whether

B has a higher or lower potential than the earth. If the student does this correctly, he will find that the potential of *B* is higher than the earth; in other words, that it has a *positive* potential. Moreover, he should be able to decide that all points on *B* are at a positive potential; for, according to paragraph 3, sect. 504, they all have the same potential. The student will thus arrive at a decision which is really a general rule:

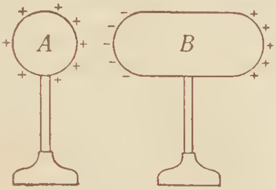


FIG. 320

Rule 1. Bringing a positive charge near an insulated conductor raises the potential of that conductor.

Take another example,—one where a *negative* charge is brought near a conductor. In this case, as a result of the force exerted by the negative charge of *A* (Fig. 321), the negative charge on *B* is repelled. If we should connect *B* to the ground, a negative charge would flow from *B* to the ground. Hence we may deduce from paragraph 2, sect. 504, that the potential of *B*, in the condition shown by Fig. 321, is less than that of the ground: its potential has been lowered. This is an example of the following general rule:

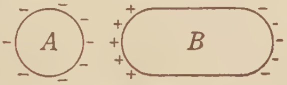


FIG. 321

Rule 2. Bringing a negative charge near an insulated conductor lowers the potential of that conductor.

It is not difficult to prove Rules 1 and 2 experimentally. If an electrostatic voltmeter or a simple leaf electroscope is connected to the conductor *B*, the changes in the potential of *B* can readily be tested when *A* is moved nearer or farther away. An electrostatic voltmeter of the common type does not distinguish between positive and negative potentials; for when it is positively charged, the same kind of deflection is produced as when it is charged by an equal negative charge. Hence in performing this experiment it is best to start *B* with a small excess of positive electricity; that is, at a positive potential when *A* is removed. If *A* is positively charged, the deflection will be increased when *A* is brought near B (Rule 1); however, if A is negatively charged, the deflection will be decreased (Rule 2).

507. A grounded conductor brought near a charged conductor. *Rule 3. Bringing a grounded conductor near an insulated charged conductor reduces the difference of potential between the charged conductor and the ground.*

This rule can either be proved by experiment or be deduced from the principles and rules already given. An experimental method will first be considered. Let A (Fig. 322) be a charged conductor connected to an electroscope or an electrostatic voltmeter (the deflection of a simple electroscope indicates the difference of potential between whatever it is connected to and its case, which ought to be grounded). When the conductor B, which is connected to the ground, is moved nearer A, the leaves converge, showing that the potential of A is lowered. If B is moved farther away, the divergence of the leaves increases and the potential of A is raised. The experiment may be repeated with A charged negatively. In this way Rule 3 can be verified.

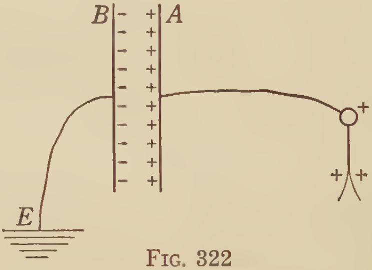


FIG. 322

A simple theoretical proof of this rule is easily obtained. When B is charged as shown in Fig. 322, moving it nearer to A is seen to be an example of Rule 2. Or, if A is negatively charged, B will have a positive charge by induction, and Rule 1 applies.

Another theoretical method is based on the definition of difference of potential (paragraph 5 of section 504). The potential of A is equal to the work done in carrying a unit positive charge from the ground to A. Work is done in this case because the charge on A repels the unit charge. When B is near, *less* work will be done; for the negative charge on B will attract the unit positive charge and partly neutralize the repellent force of the charge on A. If the work done in carrying the unit charge from the ground to A is less when B is near A, the difference of potential between A and the ground must be less.

508. The condenser. Let the two plates of Fig. 323 be connected to some source—a battery of dry cells, or an electrostatic machine, or a dynamo—giving a definite difference of potential between its terminals. As soon as *A* and *B* are connected, charges will flow into them until the difference of potential between them becomes the same as that of the source. Now if *A* and *B* are moved closer together, this motion (apply Rules 1 and 2) will decrease the *difference* of potential between the two, and more electricity will flow into the two conductors until they are again at the same difference of potential as the source. *Larger charges on A and B are required, to establish the same difference of potential, when the conductors are closer together.*

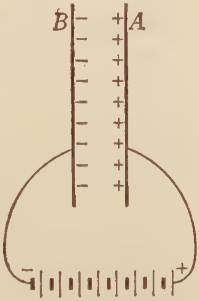


FIG. 323

As an example, assume that the plates are several centimeters apart, and that they have been charged by connecting them to a battery having a large number of cells in series and giving 1500 volts. If the plates are disconnected from the battery and are well insulated, the difference of potential between them will remain at 1500 volts. If the plates are now moved closer together, so that the distance between them is, say, half a centimeter, the difference of potential will fall, perhaps to 300 volts or lower. If they are again connected to the charging battery, more charge will flow in until the difference of potential is again 1500 volts. If the battery had not been disconnected, the charge would have flowed in while the plates were in motion. The relatively large charges on two conductors which have been brought close together are well shown by experiments with a Leyden jar (sect. 395).

Since it was believed that electricity was some sort of fluid which could be "condensed" into two plates, such a pair of plates was long ago called a *condenser*. While the original idea was erroneous, the name is still used.

A condenser may consist of two metallic plates insulated from each other, as in a Leyden jar, or of a large number of plates. In Fig. 324, *A* represents a group of five plates, and *B* another group of five, the plates being seen edge on. The plates of each group are connected together, but each group is carefully insulated from the other. In order to get them close together and yet insulated, the plates are often separated by thin sheets of mica or by paper saturated with paraffin wax.

The essential principle as just presented may be summarized as follows: (1) *Bringing two oppositely charged conductors near together tends to lower the difference of potential between them. Hence* (2) *larger charges are required in order to produce the same difference of potential.*

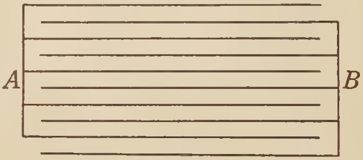


Fig. 324

509. A condenser connected to an alternating-current circuit. A very interesting illustration of the action of a condenser is its use on an alternating-current circuit. If an ordinary telephone condenser of a size marked *2 microfarads* (symbolized in Fig. 325 by the two parallel lines at *C*) is connected through a 25-watt 110-volt tungsten-filament lamp, *L*, to the alternating-current light-circuit, the lamp will glow. This is an interesting case, for there is not a complete electric circuit in the usual sense. The plates of the condenser are insulated from each other, and current cannot flow through the condenser. The explanation is not difficult and will aid one to a better understanding of the action of a condenser.

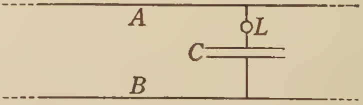


Fig. 325

In the alternating-current circuit the difference of potential, or voltage, between the wires is continually changing. At the instant when the wire *A* is charged positively, the wire *B* is charged negatively. Then the voltage between them decreases until in a very short time they are at the same potential. Later the wires become charged in the opposite direction, *A* becoming negative and *B* positive, and so on, the potential changing to and fro at a rapid rate. Knowing what is happening to the wires *A* and *B*, one can next explain why the lamp will glow. When *A* is positive and *B* negative, the condenser will become charged. When the difference of potential begins to decrease,—that is, when A begins to lose its positive charge,—the condenser begins to discharge back through the lamp. Then, as A becomes negative and B positive, the condenser charges up in the opposite direction, the charging current going through the lamp. Hence at each change in the direction of the voltage of the line—at each "oscillation"—the condenser is discharged and then charged oppositely. If the wires A and B are connected to the usual type of alternating circuit, the wire A will be charged positively 60 times a second. Hence the condenser will be fully charged and discharged 120 times a second. On each charge or discharge a current passes through the lamp and gives what we call an alternating current in the lamp.*

If the condenser is connected in a circuit which has a voltage acting always in one direction,—not pulsating nor alternating,—there will be no continual charging and discharging through the lamp, and hence the lamp will not glow.

510. Quantitative relation between the charge and the voltage of a condenser. We are now ready to discuss the quantitative relation between the *quantity* on a condenser, its *voltage*, and what is called the *capacity*, or *capacitance*, of the condenser. That quantitative relations in electricity are very important has been pointed out a number of times, and this case is no exception to the general rule.

In the case of the condenser there is a very simple relation between the charge on the condenser and the voltage between the plates; for the *charge, or quantity, is directly proportional to the voltage*. For example, if the voltage is doubled, the charge is doubled. This can be expressed algebraically by the equation

$$Q = CV, \tag{88}$$

where Q represents the charge† and V the voltage, or difference of potential, between the plates. C is a constant for any one condenser.

---

* When a 2-microfarad condenser and a 25-watt 110-volt lamp are connected to a 110-volt 60-cycle circuit, the current through the lamp is about 0.08 ampere.

† By Q, or the charge on a condenser, is meant the charge on one plate only. In symmetrical condensers the charges on the two conductors are equal, but of opposite signs.

511. Capacity, or capacitance. If the constant C of equation (88) is known for any condenser, the charge Q on the condenser can always be obtained by multiplying the constant by the voltage. Obviously, then, a knowledge of the value of the constant C for a condenser not only gives the student some idea of the size of the condenser but enables him to compute easily the value of Q. In fact, this constant is such a useful and convenient one that it is given a name—the *capacity*, or *capacitance*,* of the condenser.

The student needs a word of caution about the use of the term *capacity*. The most common use, in other than electrical practice, is in connection with the volume measurement of commodities. In such cases the word usually means the total amount the given container will hold. But it does not have that sense in electrical usage nor in some other cases. Consider the case of a tank used to hold compressed air or some other gas. If the gas obeys Boyle's law (sect. 45), the quantity of gas in the tank is given by

$$Q = kp,$$

where p is the pressure of the gas, and k is a constant which depends on the size of the tank. If one wishes at any time to know the quantity of gas in the tank, one reads a pressure gauge and then multiplies the pressure by the constant k. The product will give the quantity of gas in the tank. The constant gives one also some idea of the size of the tank, and it may be called the *capacity* of the tank. It is not, however, the total amount of gas that the tank will hold. Certainly no one would say that the capacity of the tank is the total amount of gas that the tank will hold, for that is an indeterminate quantity. The total amount of gas that it will hold can be determined only by compressing gas into the tank until it bursts.

In electrical terminology the word *capacity* is used in the sense just explained in the case of the gas tank. *It does not mean the total amount of electricity the conductor will hold.*† A little reflection will convince one that the total amount of electricity that a conductor can hold is indeterminate. One may keep adding charges until something smashes or until a spark breaks down the insulation.

---

*The American Institute of Electrical Engineers has recommended that *capacitance* be used instead of *capacity*, because the word *capacity* is often misleading.

†Similarly, the *thermal capacity* of a body is not equal to the total amount of heat that the body will hold.

The quantitative definition of the capacitance, or capacity, of a condenser may be stated in either one of two ways. Both are based on equation (88). From this equation it is readily seen that

$$C = \frac{Q}{V}. \tag{89}$$

Hence *the capacitance is equal to the ratio between the charge and the voltage of the condenser.* Another way is to take a special case —one where the voltage is unity (that is, where $V = 1$). In this case $C = Q$. Hence *the capacitance of a condenser is equal to its charge when the charging voltage is equal to unity.*

512. Capacitance dependent on physical dimensions. In section 508 an experiment was described in which moving two plates together decreased their voltage. In this case the plates were insulated, and the quantity of electricity on them remained constant. This experiment shows that moving the plates together *increases the capacitance*; for since Q was constant, V decreased, and C is equal to Q/V, it follows that C was increased. In general, moving two conductors nearer together increases the capacitance of the pair.

A simple experiment with a roll of tin foil shows that an increase in the charged area increases the capacitance. The foil is rolled on a tube or rod made of some insulating material. While rolled up, it is charged and then touched to an electroscope or electrostatic voltmeter. The size of the deflection will give some idea of the potential of the foil. If the foil is partly unrolled without discharging it, one will find by again touching it to the electroscope that its potential is now less; as we saw in the last paragraph, this means that the capacitance is greater. When the tin foil is rolled up again, it will be found to have again a larger potential and hence a smaller capacitance.

Any charged conductor (such as a roll of tin foil) may be regarded as part of a condenser, the other part being the walls, floor, table, etc., where there will be an equal and opposite induced charge. In the experiment of measuring the potential of the foil, what was really measured was the *difference* of potential between the foil, on the one hand, and the walls, floor, table, etc., on the other.

513. Capacitance dependent on the medium between the plates. If instead of using air as the insulator between the two plates of a condenser, one uses a sheet of glass, hard rubber, or some other insulating medium, the capacitance of the condenser becomes larger.

This is easily shown experimentally. In Fig. 326 the plate A is connected to an electroscope or electrostatic voltmeter and then charged. The plate B is connected to the ground. When a plate of glass or some other insulating material is slipped between the two plates, the leaves of the electroscope will fall. When the plate of insulating material is withdrawn, the leaves will again diverge. The electroscope thus shows that the *potential* of A is *decreased* by the insertion of the glass plate. As before, a lowering of the difference of potential, with the *charge Q remaining constant*, means, since $C = Q/V$, that the capacitance has become larger.

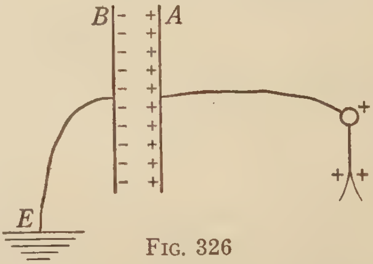


FIG. 326

The experiment showing the change in the capacitance and potential when the insulating medium between the conductors is changed is a very important one. It shows that we are concerned with something more than with one charge acting across space on another charge. It shows that the medium between the conductors has something to do with the electrical action. Indeed, from this experiment it is possible to deduce the fact that two charges of electricity, when submerged in some insulating oil, should attract each other with a very different force from that exerted in air. It was such experiments as this that suggested the "strain" theory referred to in section 393. It led physicists to think more about the medium in which electric and magnetic fields are produced. It ultimately led to the discovery of electromagnetic waves, which today are so commonly used in wireless telegraphy and telephony.

The medium between the conductors of a condenser is usually called the *dielectric*. Different dielectrics do not behave alike: some change the potential and capacitance more than others. The *dielectric constant k*, which is used to measure the relative effects of the different dielectrics, is defined as follows: *The dielectric constant* of any substance is equal to the ratio of the capacitance of a condenser when that substance is used as the dielectric to the capacitance when there is a vacuum between the conductors* (for all practical purposes air at ordinary pressures may be used instead of a vacuum). The numerical values of the dielectric constants of a number of substances are given on the next page.

* Sometimes called *specific inductive capacity.*

DIELECTRIC CONSTANTS

Air	1.0006	Paraffin wax	2–2.3
Ebonite	2.7–3	Oil (castor)	4.6–4.8
Glass (crown)	5–7	Oil (petroleum)	2–2.2
Glass (flint)	7–10	Shellac	3–3.7
Mica	5.7–7	Vaseline	2.2

In the construction of Leyden jars great care should be used in selecting glass which has a large dielectric constant. There is such a wide variation in the different kinds of glass that a considerable gain in capacitance is obtained by selecting the best.

The important discovery that the effect of one charge on another depends on the character of the medium between the two was first made in 1773, by Cavendish; but this extremely important discovery was lost to the world through his failure to publish it. It was only after his manuscripts were published by Maxwell, in 1879, that it became known what important discoveries had remained hidden in his papers. Meanwhile Michael Faraday, in 1837, rediscovered the effect of the medium on the capacitance of conductors and made measurements of dielectric constants. Apparently Cavendish was the first to have a clear idea of the capacitance of a conductor and actually to measure it.

In Chapter XXVI, sect. 390, it was stated that the law of force given in equation (8) is true only when the surrounding medium is air. When the medium is one having a dielectric constant equal to k, this law becomes

$$F = \frac{q_1 q_2}{kd^2}. \tag{90}$$

514. Units and formulas for capacitance. In the practical system of units the unit for capacitance is the *farad* (named after Faraday), but it is so large a unit that the millionth part of it—the *microfarad*—is commonly used. For very small capacitances, such as those used in the receiving circuits of wireless telephony, a still smaller unit (one millionth of a microfarad), the *micromicrofarad*, is used. The capacitance of a Leyden jar is a small fraction of a microfarad. For example, a common type of jar, $10\frac{3}{8}$ inches high and $4\frac{1}{4}$ inches in outside diameter, with the tin foil extending up $7\frac{1}{8}$ inches from the bottom, has a capacitance of one six-hundredth of a microfarad.

The electrostatic unit of capacitance is a much smaller unit than a microfarad.

1 microfarad = 900,000 C.G.S. electrostatic units.

The capacitance of an isolated sphere, in electrostatic units, is numerically equal to the radius of the sphere measured in centimeters.

The capacitance of two parallel plates, in electrostatic units, is (provided the plates are large compared with their distance apart)

$$C = \frac{Ak}{4\pi d}, \tag{91}$$

where A is the area of one side of one of the plates, k the dielectric constant, and d the distance between the plates.

In many other cases where the surfaces are simple geometrical figures it is possible to compute the capacitance from the dimensions. The advanced student, however, will learn experimentally some of the laboratory methods of determining capacitance.

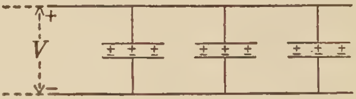


FIG. 327. Condensers in parallel

The fundamental equation of capacitance ($Q = CV$) holds for all systems of units, provided that one is consistent in using the same system for all three quantities. In the practical system

Quantity in coulombs = capacitance in farads × voltage.

515. Condensers in parallel and in series. Connecting three condensers in *parallel*, as shown in Fig. 327, *increases* the capacitance. The resulting capacitance for parallel arrangement is *the sum of the separate capacitances*. The formula is

$$C = C_1 + C_2 + C_3, \tag{92}$$

where C is the total capacitance, and C_1, C_2, and C_3 are the capacitances of the separate condensers.

This law is easily proved. When the condensers are connected in this way, each condenser is charged up to the same voltage, V. The charge on the first condenser is

$$Q_1 = C_1 V; \tag{88 a}$$

on the second,

$$Q_2 = C_2 V; \tag{88 b}$$

and on the third,

$$Q_3 = C_3 V. \tag{88 c}$$

The total charge, Q, in the system (that is, the total charge obtained when the combination is discharged) is

$$Q = Q_1 + Q_2 + Q_3.$$

From equations (88 *a*), (88 *b*), and (88 *c*) this becomes

$$\begin{aligned} Q &= C_1V + C_2V + C_3V \\ &= (C_1 + C_2 + C_3)V. \end{aligned} \qquad (93)$$

If C is the total capacitance and Q the total charge, we have, from the fundamental equation of capacitance (equation (88)),

$$Q = CV.$$

By comparing equations (93) and (88) we get equation (92), which we were to prove.

If there are more than three condensers, more terms must be added. In general, we may say that

$$C = C_1 + C_2 + C_3 + C_4 + \cdots. \qquad (94)$$

When condensers are connected in *series* (Fig. 328), the effect is to decrease the capacitance. The resulting capacitance is less than that of any one of the condensers. When they are connected in this way, no one condenser is charged up to the full voltage, V; hence each has a smaller charge than if it were directly connected to the terminals of voltage V. The formula for the series combination is

$$\frac{1}{C} = \frac{1}{C_1} + \frac{1}{C_2} + \frac{1}{C_3}, \qquad (95)$$

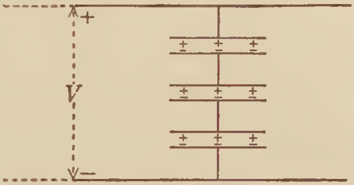


FIG. 328. Condensers in series

where C is the resulting capacitance, and C_1, C_2, and C_3 are the capacitances of the separate condensers. This formula will now be derived.

If Q_1, V_1, and C_1 are the quantity, voltage, and capacitance of the first condenser, we have

$$Q_1 = C_1V_1. \qquad (96)$$

Similarly, for the second,

$$Q_2 = C_2V_2, \qquad (97)$$

and, for the third,

$$Q_3 = C_3V_3. \qquad (98)$$

The total difference of potential across all the condensers is the sum of that across each condenser, or

$$V = V_1 + V_2 + V_3.$$

If we substitute in this equation the values of V_1, V_2, and V_3 obtained from equations (96), (97), and (98), we get

$$V = \frac{Q_1}{C_1} + \frac{Q_2}{C_2} + \frac{Q_3}{C_3}. \tag{99}$$

An inspection of the method of connecting the condensers will show that the charge on one condenser must be equal to that on the next. With the exception of the first and the last, the plates are insulated from the charging terminals, and the charging is merely a separation of the positive and negative charges by induction. When separation takes place by induction only, the two charges are equal in quantity. Hence

$$Q_1 = Q_2 = Q_3.$$

When the series combination is discharged, the charge obtained is that of the outside plates, the interior charges simply uniting. Hence we have for the charge, Q, obtained in the discharge,

$$Q = Q_1 = Q_2 = Q_3.$$

Substituting Q for Q_1, Q_2, and Q_3 in equation (99),

$$\begin{aligned} V &= \frac{Q}{C_1} + \frac{Q}{C_2} + \frac{Q}{C_3} \\ &= Q\left(\frac{1}{C_1} + \frac{1}{C_2} + \frac{1}{C_3}\right). \end{aligned} \tag{100}$$

From the fundamental equation, (88),

$$V = Q\frac{1}{C}, \tag{101}$$

where C is the total capacitance. Comparing equations (100) and (101), we see that (equation (95))

$$\frac{1}{C} = \frac{1}{C_1} + \frac{1}{C_2} + \frac{1}{C_3}.$$

In case there are more than three condensers, it can be shown that other terms should be added to equation (95). For example, for four condensers,

$$\frac{1}{C} = \frac{1}{C_1} + \frac{1}{C_2} + \frac{1}{C_3} + \frac{1}{C_4}.$$

If two equal capacitances are connected in series (that is, when $C_1 = C_2$),

$$\frac{1}{C} = \frac{1}{C_1} + \frac{1}{C_2} = \frac{2}{C_1},$$

or

$$C = \tfrac{1}{2} C_1.$$

In a similar way it can be shown that when three equal capacitances are connected in series, the resultant capacitance is one third of one of them.

Commercial condensers are usually made of a number of condensers connected in parallel (see Fig. 324). If the condenser is to be used on high potentials, it may be made of a number of plates of metal (often sheet copper) with glass plates between, the whole being immersed in oil. Small condensers made of sheets of tin foil separated by paraffined paper are used on low voltages. By using tin foil and thin paper very compact condensers are obtained. Some telephone condensers, having a capacitance equal to that of several hundred Leyden jars, are so compact that one can be put in a vest pocket.

Electric cables, such as those placed under the ground or in the sea, have capacitance, and behave like condensers. The capacitance effect of these is often extremely important. If a line for transmitting electric power has two wires, one will be charged positively and the other negatively, and together they will form a charged condenser. When the line is long, the capacitance becomes an important factor. It is because of the practical effects of capacitance in commercial work that one interested in practical electricity should thoroughly understand the subject matter contained in this chapter.

516. Energy of a charged condenser. When a charge of electricity, Q, is carried from a region where the potential is zero to a place where the potential is V, the work done is equal to QV (see equation (18), sect. 416). By a similar rule, when water of weight w is carried from zero level up to a level l, the work done is wl. In some cases these simple expressions do not apply. Take, for example, the case of filling a reservoir. If the bottom of the reservoir is at zero level and the top of the water is at a level l, the work done in filling the reservoir is

$$W = \tfrac{1}{2} wl.$$

This expression represents the weight of water times the *average* height it has been lifted. Similarly, when a condenser is charged, the potential difference between the plates is gradually raised from zero to the final potential, V, and the work done in charging it is

$$W = \tfrac{1}{2}QV, \tag{102}$$

where Q is the final charge. Since the energy of a charged condenser must be equal to the work done in charging it, this expression gives the energy possessed by the condenser.

If the charge in equation (102) is measured in *coulombs*, and the potential difference in *volts*, the energy will be in *joules*. If both Q and V are measured in C.G.S. units, the energy will be in ergs.

Since any charged conductor may be regarded as a condenser, the other "coating" being the surrounding objects (walls, ceiling, table, etc.), equation (102), for the energy of a charged condenser, applies to any charged conductor.

517. Summary. It is customary to take the potential of the earth as zero and to call all potentials higher than this *positive*, and all lower potentials *negative*. When the potential of the earth is taken as zero, the expression "the potential of a conductor" really means the *difference of potential* between that conductor and the ground.

Rule 1. Bringing a positive charge near an insulated conductor raises the potential of that conductor.

Rule 2. Bringing a negative charge near an insulated conductor lowers the potential of that conductor.

Rule 3. Bringing a grounded conductor near an insulated charged conductor reduces the difference of potential between the charged conductor and the ground.

When two oppositely charged conductors are moved closer together, their difference of potential will be decreased if they are insulated. However, if they are connected to some source which maintains a constant difference of potential, the quantity, or charge, on the conductors will increase as they are moved nearer each other. When the conductors are close together, they will have, for the same difference of potential, relatively large charges. Such a pair of conductors is called a condenser. A condenser may also consist of two groups of plates placed near together. A Leyden jar is an example of one type of condenser. As these jars may be charged to high potentials, relatively large charges may be stored in them.

When a condenser is connected in an alternating-current circuit, there will be a current due to the charging, discharging, and charging in an opposite direction, which follow one another at a rapid rate. When a condenser is connected in a direct-current circuit, there is no continuous current in the connecting wires.

The equation $Q = CV$ (equation (88)) gives the quantitative relation between the charge on the condenser and the difference of potential, V. C, called the *capacitance* or *capacity*, is a constant for any one condenser.

The capacitance of a condenser is equal to the ratio of the charge on the condenser to the difference of potential; that is,

$$C = \frac{Q}{V}.$$

In other words, the capacitance of a condenser is equal to the charge when the voltage is equal to unity.

The capacitance of a condenser depends on the distance between the conductors: the less this distance, the greater the capacitance. It depends also on the size of the conductors. A very important discovery, first made by Cavendish, and independently by Faraday, showed that the capacitance of two conductors depends on the *medium* between them. From this it can be deduced that the attraction and repulsion of charges are different for different media.

The insulating medium surrounding any conductor is called the *dielectric*. The dielectric constant of any material is equal to the ratio of the capacitance of a condenser when that material is used as the dielectric, to the capacitance when there is a vacuum between the conductors.

When two charges, q_1 and q_2, are on two conductors the sizes of which are small compared with their distance apart, the law of force becomes equation (90),

$$F = \frac{q_1 q_2}{kd^2},$$

where d is the distance apart, and k is the dielectric constant of the medium in which the charges are immersed.

When Q is measured in coulombs and V in volts, the capacitance, C, in equation (88) is measured in *farads*. The farad is such a large unit that the microfarad is commonly used.

The C.G.S. electrostatic unit of capacitance is relatively small, for

$$1 \text{ microfarad} = 9 \times 10^5 \text{ electrostatic units.}$$

The capacitance in electrostatic units of a sphere which is surrounded by air, but isolated from other conductors, is equal to the radius of the sphere expressed in centimeters. The capacitance of two parallel plates d centimeters apart (provided the plates are large compared with d) is (equation (91)) $Ak/4\pi d$ electrostatic units, where A is the area of one side of one of the plates, and k is the dielectric constant of the medium between the plates.

When the capacitances C_1, C_2, C_3, C_4, etc. are connected in parallel, the resultant capacitance is (equation (94))

$$C = C_1 + C_2 + C_3 + C_4 + \cdots.$$

When the capacitances C_1, C_2, C_3, C_4, etc. are connected in series, the resultant capacitance is obtained from

$$\frac{1}{C} = \frac{1}{C_1} + \frac{1}{C_2} + \frac{1}{C_3} + \frac{1}{C_4} + \cdots.$$

The energy of a charged condenser is given by equation (102),

$$W = \tfrac{1}{2}QV.$$

If Q is measured in coulombs and V in volts, the energy, as computed from this equation, will be expressed in joules. If both Q and V are in C.G.S. units, the energy is given in ergs. This expression for the energy may be applied to any charged conductor if V is regarded as the difference of potential between it and the ground.

PROBLEMS

1. Find the capacitance (*a*) in electrostatic units and (*b*) in microfarads of a sphere 60 cm. in diameter.

2. The radius of the earth is approximately 6400 km. What is its capacitance in microfarads?

3. A pair of parallel circular plates are 30 cm. in diameter and 0.5 cm. apart, with air between them. Compute the capacitance (*a*) in electrostatic units; (*b*) in microfarads.

4. What would be the capacitance, in microfarads, of the pair of plates of Problem 3 if they were submerged in oil the dielectric constant of which is 2.0?

5. The capacitance of a small plate condenser was found by experiment to be 0.00015 microfarad in air, and 0.00033 microfarad when submerged in engine oil. What is the dielectric constant of the oil?

6. Each plate of the condenser of Fig. 324 was 25 × 30 cm., and they were separated by 3 mm. of glass the dielectric constant of which was 6.0. Find the total capacitance in microfarads.

7. Three condensers, of capacitance 0.5, 1.0, and 2.0 microfarads, respectively, are connected in series. (*a*) Find the resulting capacitance. (*b*) What will be the charge when a voltage of 500 volts is applied to the combination?

8. Three condensers, of capacitance 0.5, 1.5, and 2.0 microfarads, respectively, are connected in parallel to a 220-volt circuit. (*a*) Compute the resulting capacitance. (*b*) Compute the charge in coulombs on each condenser.

9. The capacitance of each of ten Leyden jars is 0.0016 microfarad. They are connected in parallel and charged to a potential difference of 30,000 volts. Compute the total energy in joules.

10. Two charged balls attract each other with a force of 50 dynes. If they were submerged in oil without discharging them, what would be the attractive force? (The dielectric constant of the oil is 3.1.)

11. A condenser of 2.0 microfarads was charged to a difference of potential of 200 volts and then insulated. An uncharged condenser of 1.0 microfarad was then connected in parallel with the first. What is the voltage at their terminals?

CHAPTER XXXVI

ELECTRICAL UNITS

Introduction, 518. The C.G.S. electrostatic system, 519. The C.G.S. electromagnetic system, 520. The numerical relation between electromagnetic and electrostatic units, 521. The practical system, 522. International units, 523.

518. Introduction. Even a very elementary knowledge of electricity is incomplete without a knowledge of the principal electrical units and the logical processes on which they are built.

The units used in electricity beautifully illustrate the advantages of a logical system. In the C.G.S. system there is an entire absence of *arbitrary* units: they are all based on the units of mechanics—the centimeter, the second, the dyne, and the erg. It is a derived system of units. To appreciate such a system one has only to compare it with the British system of weights and measures, with its large number of arbitrary units. The practical advantage in a derived system lies in the ease with which computations can be made. For example, *watts* are obtained by multiplying *amperes* by *volts*, no numerical constant appearing in the equation. The same is true of many of the other electrical equations. Again, in changing from one unit to another in measuring the same electrical quantity, all that is usually necessary is to shift the decimal point. Contrast this with the calculations necessary in changing from one to the other of two units of volume—from cubic inches to bushels, for example, or from cubic inches to cubic yards.

There are two fundamental systems of units in electrical work: the C.G.S. *electrostatic* and the C.G.S. *electromagnetic*. There is also the *practical* system, which is based on the electromagnetic. There is a very good reason for the use of all these systems. Certain types of problems can often be solved more easily by using one system than by using another. For example, it is often simpler in dealing with certain problems involving stationary charges, such as are found in a condenser, to use C.G.S. electrostatic units; but in the great majority of cases either the C.G.S. electromagnetic or the practical system, which is a modification of the electromagnetic system, is used.

519. The C. G. S. electrostatic system. This system is based on the attraction between two stationary charges, or electrostatic forces. The *fundamental* unit is the unit of *charge*, or *quantity*. It is defined from equation (8) of section 390,

$$F = \frac{q_1 q_2}{d^2},$$

which gives the force in *dynes* between two small charges d centimeters apart in a vacuum.* From this unit and the C.G.S. units of mechanics, the other electrostatic units are derived by the aid of the fundamental equations. For example, a unit difference of potential is a difference of potential such that the work done in carrying an electrostatic unit of charge between the two places is equal to 1 erg.

520. The C. G. S. electromagnetic system. This system is based on a magnetic quantity—the *unit pole*. The first electrical term to be defined is *current*, which is defined in terms of the force, in dynes, that it will exert on a unit pole. The definition of *current* is thus based on the fact that a current produces a magnetic field; hence the name *electromagnetic*. The other units are all defined in a logical way by means of the fundamental equations of electricity and those of mechanics. This will be outlined for the more commonly used quantities. The student should note that the first is defined from its mechanical action, by a force measured in dynes. The other quantities are all defined in terms of those which precede. There is no "reasoning in a circle."

Unit pole: based on the force one pole exerts on another.
Strength of magnetic field: equal to the force the field exerts on a unit pole.
Current: based on the strength of the magnetic field produced by the current.
Quantity: the product of average current times the time of flow.
Difference of potential: work done when the unit quantity is moved.
Resistance: ratio between difference of potential and current (Ohm's law).
Capacity: ratio of quantity to difference of potential.

---

*The difference in the force between the charges when in a vacuum or in the air is usually negligible.

521. The numerical relation between electromagnetic and electrostatic units. It is found that there is in each case a remarkable relation between the corresponding units in the electromagnetic and electrostatic systems. The ratio between any two is equal to the numerical value of the speed of light (3×10^{10} centimeters per second) or in some cases to the square of it. The following table gives the values of this relation:

Current: 1 electromagnetic unit $= 3 \times 10^{10}$ electrostatic units.
Quantity: 1 electromagnetic unit $= 3 \times 10^{10}$ electrostatic units.
Difference of potential: 1 electromagnetic unit $= \frac{1}{3 \times 10^{10}}$ electrostatic unit.
Resistance: 1 electromagnetic unit $= 1/(3 \times 10^{10})^2$ electrostatic unit.
Capacity: 1 electromagnetic unit $= (3 \times 10^{10})^2$ electrostatic units.

522. The practical system. The practical system is based on the electromagnetic, and differs from it only in that units of more convenient size are used. But in making this change two things were kept in mind: first, the change was always made by some factor of 10 (that is, by merely moving the decimal point); second, the changes were so made that the fundamental equations would still hold. For example, Ohm's law is

$$I = \frac{V}{R}.$$

A unit called the *volt*, 10^8 times as large as the electromagnetic C.G.S. unit was taken; the unit thus obtained was nearly the same as the electromotive force of the Daniell cell,* which had become a sort of standard. Similarly, by taking a unit called the *ohm*, 10^9 times as large as the electromagnetic unit, a unit was obtained nearly equal to the resistance of a column of mercury 1 square millimeter in section and 100 centimeters long, which had been used as a standard of resistance. Thus two units, the *volt* and the *ohm*, were nearly equal to those which were already in use. In order that Ohm's law should still be true it became necessary to take a new unit of current, the *ampere*, which would be one tenth

*A cell of the same chemical structure as a gravity cell.

of the electromagnetic unit. This may be verified; for if $I = V/R$, then

$$\frac{I}{10} = \frac{V \times 10^8}{R \times 10^9}.$$

The following table gives the relations between the *practical* and the *electromagnetic units.* Since the practical units were defined in terms of the electromagnetic units, this table really gives the definition of the practical units.

Current: 1 ampere = 1/10 C.G.S. electromagnetic unit.
Quantity: 1 coulomb = 1/10 C.G.S. electromagnetic unit.
Difference of potential: 1 volt = 10^8 C.G.S. electromagnetic units.
Resistance: 1 ohm = 10^9 C.G.S. electromagnetic units.
Capacitance: 1 farad = $1/10^9$ C.G.S. electromagnetic unit.
Power: 1 watt = 10^7 C.G.S. electromagnetic units.
Work: 1 joule = 10^7 C.G.S. electromagnetic units (ergs).
Self-inductance and mutual inductance: 1 henry = 10^9 C.G.S. electromagnetic units.

523. International units. Since the units of the electrostatic and electromagnetic systems are derived units, it was necessary to find the fundamental quantities by *experiment.* The values of the ampere, ohm, volt, etc. were all determined experimentally. In the early days the experimental methods used in finding these quantities were not accurate as compared with the work that can now be performed in our best laboratories. As experimental methods were improved, the values of electrical units became known more and more accurately. At several different times the most recent experimental values were adopted as standard. For obvious reasons there soon arose a demand for fixed standards that would not have to be changed from time to time. Hence, in 1908, an International Electrical Conference decided to take the best-known values of the electrical standards and make *fixed* standards of them. It was recognized that these might be different from the true values, and that subsequent experiments might reveal this; however, it was decided, in the interests of stability, to ignore this, especially as it was certain that the deviations would be very small. These units were named the *international units.* The principal international units are defined as follows:

1. The *international ampere* is the unvarying electric current which, when passed through a solution of silver nitrate in water (made in accordance with certain specifications), deposits silver at the rate of 0.00111800 gram per second.

2. The *international ohm* is the resistance offered to an unvarying electric current by a column of mercury at the temperature of melting ice, 14.4521 grams in mass, of a constant cross-sectional area, and of a length of 106.300 centimeters.

3. The *international volt* is the electrical pressure which, when steadily applied to a conductor the resistance of which is 1 international ohm, will produce a current of 1 international ampere. For practical purposes the working standard for measuring voltage is the Weston cell, the electromotive force of which at 20° C. is 1.0183 international volts.

Theoretically there are two different amperes: the *international*, and the *true*, which is defined in terms of the C. G. S. electromagnetic unit. The same classification holds for the other practical units. However, the numerical difference between the two amperes, or the two ohms etc., is very small,—probably not greater than 1 or 2 parts in 10,000.

CHAPTER XXXVII

CONDUCTION OF ELECTRICITY THROUGH GASES

Conduction through gases at ordinary pressures, 524. Ionization by radiation, 525. Ionization by collision, 526. Spark discharge, 527. Discharge at reduced pressures, 528. Geissler tubes, 529. Cathode rays, 530. Nature of cathode rays, 531. X-rays, 532. X-ray tubes, 533. Nature of X-rays, 534. Electronic emission used in the rectification of alternating currents, 535. Electronic emission used in amplifiers, 536. Emission of electrons, 537. Millikan's method for measurement of the fundamental charge of electricity, 538.

524. Conduction through gases at ordinary pressures. The fact that a good electroscope will retain its charge for a long time shows that the air around it is a very poor conductor. All gases at ordinary temperatures and pressures are usually good insulators. But there are circumstances under which air becomes conducting; and when it does, it seems to behave in a manner somewhat similar to that of an electrolyte. It will be remembered that an electrolyte is a conductor because it is *ionized*; that is, it contains charged *ions* which are free to move. To make a gas a conductor some of its molecules or atoms must be separated into oppositely charged parts, or ions. These free ions may be electrons, or charged atoms, or charged clusters of atoms.

Chemical changes may leave a gas temporarily ionized; for example, the vapors arising from a flame are usually ionized for a short time, often less than a second. When a candle flame is held underneath a charged conductor, the conductor will be quickly discharged; or the candle flame may be held somewhere near and the rising vapors blown against the conductor. It is very easy to show in this way that vapors may conduct electricity. Other methods by which gases at ordinary pressures may be made conducting will be discussed later.

525. Ionization by radiation. There are several forms of radiation that will ionize any gas they pass through. The short light-waves, called *ultraviolet* light, will ionize air slightly. The still shorter waves, X-rays (Röntgen rays), produce an effect easily detected. If an X-ray tube is operated within fifteen or twenty feet of a charged electroscope, the electroscope will lose its charge quickly because the surrounding air becomes a conductor. Any metal screen that will cut off the X-rays from the neighborhood of the electroscope will destroy the effect. The radiations from radioactive substances produce somewhat similar effects.

526. Ionization by collision. It is now generally believed that when a molecule is hit by a swiftly moving electron or by a high-speed atom, one of the usual results of the collision is a separation of the molecule into oppositely charged parts, or ions.

527. Spark discharge. It is well known that when a sufficiently high potential is used, a gas becomes conducting, and a spark, or arc, discharge takes place. This will be explained by the "ionization by collision" theory.

A gas is never entirely free from ions. Between the terminals of a spark gap there will always be found a few charged ions. As soon as the terminals are charged, these ions move toward the charged terminals. If the charge on the terminal is large,—that is, if the voltage across the spark gap is large,—the ions will move rapidly. With sufficiently large attracting charges on the terminals, they will move fast enough to produce other ions from the molecules they hit. These newly formed ions at once begin to move toward the terminals and, gaining speed, will quickly have enough energy to ionize other molecules with which they collide. These in turn will produce others, so that many ions are quickly formed, and a heavy discharge can take place.

528. Discharge at reduced pressures. If a long glass tube with metallic terminals sealed in the ends is connected to the secondary of an induction coil while the air is being pumped out, the remarkable changes that take place in the appearance of the electrical discharge through the tube can be observed. One can form definite ideas of the appearance only by actual observation.

The first discharge through the tube, as the pressure is being reduced, will resemble a spark discharge in air, but the line of the discharge gradually broadens out until it fills the whole tube.

Usually the residual gas will glow with a brilliant color, which will change as the pressure changes. The colors will be different for different kinds of gases. The two ends of the tube will not have the same appearance. If air is used, there will be a blue glow around the negative electrode, while around the positive terminal there will be a rose-colored glow which will extend out along the tube almost to the blue glow of the cathode. This rose-colored glow, called the positive column, is often not continuous but shows alternate dark and bright layers called *striations*. As exhaustion proceeds, there appears next to the cathode a narrow dark region called the *cathode dark space* or the *Crookes dark space*. At lower pressures this dark space widens out, and both the blue glow and the positive column recede, gradually losing their brilliance.

When the cathode dark space extends a centimeter or so from the cathode, the glass walls of the tube in the neighborhood will begin to glow with a greenish-yellow color if the tube is made of the usual soft soda glass. When a rather high degree of exhaustion has been reached, this yellow glow may extend over the entire surface of the tube.

The conductivity of the air in the tube reaches a maximum at a pressure of about one or two millimeters of mercury. At lower pressures the conductivity gets less, and it requires a higher potential to start a discharge. It is possible, with prolonged pumping and by using the best methods of exhausting the air, so to reduce the pressure in the tube that a discharge will no longer pass through it.

It is now generally agreed that at ordinary pressures ions in the air are usually charged molecules. As the pressure is decreased, the electrons take a more and more important part in the conduction. At low pressure practically all the negative ions are electrons, while the positive ions are atoms which have lost one or more electrons.

The conductivity of the residual gas in these tubes is undoubtedly caused by collision. The few stray ions present in the gas will begin moving when the electrodes are charged. If the voltage is high enough, these ions will move fast enough to form others by collision, these newly formed ones will form others, and so on. The greater conductivity of the gas when it is rarefied is probably due to the fact that the ions, on the average, move farther before colliding with a molecule, and hence they can more easily acquire a speed sufficient to produce other ions by collision.

529. Geissler tubes. Geissler tubes are tubes that have been exhausted to somewhere near the maximum conductivity of the gas, or about to the point where the tube glows with its greatest brilliance when an electric current passes through it. The color of the light emitted depends on the kind of gas contained in the tube. The spectrum (Chapter XLV) of this light is of such a character that these tubes are very useful in advanced work in optics.

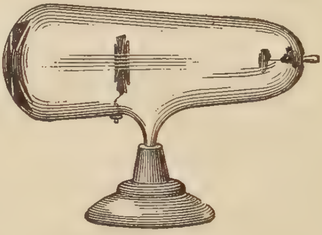


FIG. 329

530. Cathode rays. The discharge in tubes that have been exhausted rather highly shows some very remarkable properties, which have led to some very important discoveries. When the tube is highly exhausted, there is emitted from the neighborhood of the cathode a radiation known as *cathode rays*. The properties of these rays can be best understood by a description of certain important experiments.

1. *Cathode rays produce fluorescence.** The yellow glow on the walls of the tube has been referred to in section 528. When a tube such as that shown in Fig. 329 is used, it becomes obvious that the cause of this glow, or fluorescence, of the glass is an emission from the cathode; for the cross casts a shadow on the glass, which moves when the cross is shifted. Even a casual inspection of this shadow shows that the source of the emission is the cathode. When almost any substance is placed in front of the cathode in a highly exhausted tube, it will become fluorescent when struck by this radiation. Some substances fluoresce brilliantly under the action of cathode rays.

2. *Cathode rays travel in straight lines.* This conclusion is reached by an inspection of the shadows cast in a tube such as the

---

* *Fluorescence* is the name commonly applied to this phenomenon, but, strictly speaking, *luminescence* is the proper term.

one shown in Fig. 329, or by the use of the tube described in paragraphs 3 or 4 below.

3. *Cathode rays are deflected by a magnetic field.* If a screen of thin aluminum or mica with a narrow slit cut in it is placed in front of a cathode, a narrow beam of cathode rays can be sent lengthwise through the tube. The track of these rays can be made luminous by placing lengthwise in the tube a long screen covered with some fluorescent chemical. This screen is slightly inclined, so that the narrow beam of rays strikes it at almost grazing incidence over a relatively long distance. Such a tube is shown in Fig. 330. If a pole of a magnet is brought near the path of the rays, they will be deflected in a direction at right angles to the magnetic field.

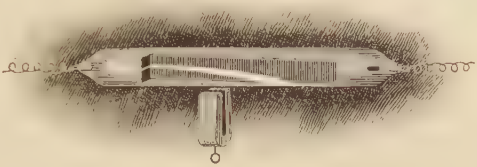


FIG. 330

4. *Cathode rays are deflected by an electric field.* In Fig. 331, A is a diaphragm with a small hole in it. The cathode rays, starting from the cathode (which should be on the left of A), are cut down to a narrow beam which passes between the two parallel metallic plates BB', and which, on striking the screen S, produces a brilliant luminous spot. If the two plates B and B' are now charged, one positively and the other negatively, the beam, as shown by the motion of the luminous spot, will be deflected *toward the positively charged plate.* (Obviously the screens and plates of Fig. 331 must be inclosed in an evacuated tube with proper terminals.)

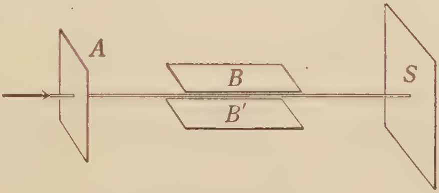


FIG. 331

5. *Cathode rays can be focused.* The simplest way of doing this is to use a concave cathode. When such a cathode is used, the rays converge toward a focus, as shown in Fig. 332.

6. *Cathode rays carry negative charges of electricity.* Carefully

conducted experiments show that when cathode rays impinge on a target, they tend to charge the target with negative electricity. This experiment is not always easy to show, but there is no doubt now that negative electricity is carried from the cathode with the rays.

7. *Cathode rays produce heat on any target they strike.* This is best shown by a simple demonstration tube which has been constructed for this purpose. In this tube a concave cathode is used which will bring the rays to a focus at a point. At this focus a piece of thin platinum foil is mounted. When the discharge is sent through the tube, the platinum foil becomes heated white-hot at the focal point. If a powerful induction coil is used to produce the difference of potential between the terminals of the tube, a hole can easily be melted in the foil.

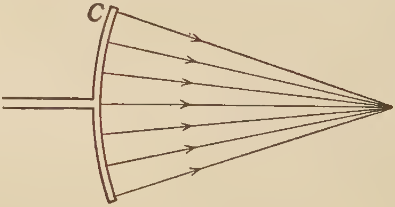


FIG. 332

8. *Cathode rays travel at high speeds.* The experiments by means of which the velocity of cathode rays was actually measured are too involved to be explained here; but the results showed that the speed depends on the voltage impressed on the tube, and that the speed varies from one tenth to about one half of the speed of light. (The speed of light is 300,000,000 meters per second.)

9. *Cathode rays will penetrate thin sheets of aluminum foil.* When screens of thin aluminum foil are used, cathode rays sufficient in quantity to excite a fluorescence of a suitable substance will penetrate the aluminum.

10. *Cathode rays produce X-rays.* This will be explained later, in sections 532–534.

531. Nature of cathode rays. There are only two forms of radiation known:

1. *Radiation by wave-motion*; for example, the propagation of tremors, or waves, through a bridge or a building. Sound and light are examples of this sort of radiation.

2. *The transfer of energy by material particles*; for example, by a stream of machine-gun bullets.

Cathode rays have three properties that prohibit their classification under wave-motion. These are discussed in the last section, paragraphs 3, 4, and 6. There are no reasons known why either an electric field or a magnetic field should deflect a train of waves. As regards paragraph 6, all experiments indicate that electricity can be carried only by the motion of some form of matter. Hence, when it is found that cathode rays always carry negative electricity, we are warranted in thinking that a beam of cathode rays is a stream of negatively charged particles. If this is true, the fact that an electric field deflects the stream is at once explained. As regards the magnetic deflection (sect. 530, paragraph 3), it must be remembered that a stream of negatively charged particles is an electric *current*. Since a current has a current-magnetic-field force exerted on it when it flows through a magnetic field, we should expect a stream of charged particles to be deflected by a magnet. When the direction of a current and that of the magnetic field are known, the direction of the deflecting force can be predicted. The observed deflections of the cathode rays are in the predicted direction if the rays are a stream of negatively charged particles.

There remains the interesting question, What are these charged particles which constitute the cathode rays? Sir William Crookes, who was one of the first to devote much attention to the study of cathode rays, thought that they were molecules. Later, experiments were devised that led to an estimation not only of the *speed* of these charged particles but also of the *mass*. It was found that the mass of one of them was equal to about one eighteen-hundredth of the mass of the hydrogen atom, the lightest of all known atoms. This identifies these particles with the *electrons*, to which reference has been made so often; in fact, this was the first discovery of the existence of electrons.

Hence it is now regarded as proved that *cathode rays consist of a stream of large numbers of electrons moving at high speeds.*

532. X-rays. Whenever cathode rays strike any substance, they cause that substance to emit a radiation known as *X-rays* (sometimes called *Röntgen rays,* after their discoverer).

The X-rays emitted by the heavier metals, such as platinum and tungsten, are very penetrating. As is commonly known, X-rays have been of tremendous aid to the physician and the surgeon. As a very simple example, suppose that one holds the hand in a beam of X-rays which falls upon a screen and causes it to fluoresce. As the fleshy parts of the hand are relatively transparent, while the bones, being denser, are more opaque, a "shadow" of the hand, showing the bones, will be seen on the screen. If this shadow falls on a sensitive photographic plate, a permanent register is obtained when the plate is developed and fixed; for X-rays have an actinic effect, similar to that of light, on the sensitive emulsion of a photographic plate (see Fig. 333). X-rays cannot be seen directly by the eye, but the fluorescence produced on a screen by the rays can be seen. A fluorescent screen may be used in diagnosis, but it is not so satisfactory as a permanent record, which is obtained on a photographic plate.

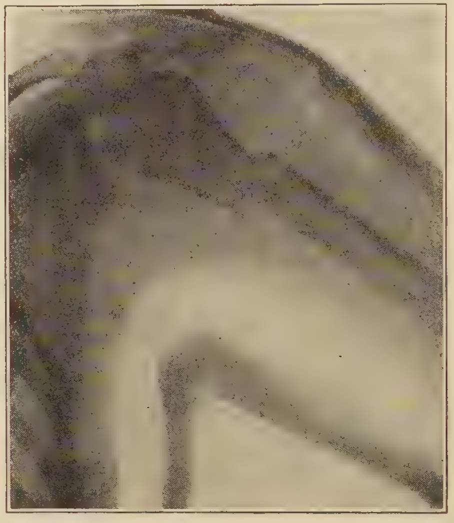


FIG. 333. An X-ray photograph of an arm broken near the shoulder

It is only where there are objects of different densities placed in the path of the rays that any difference is shown by X-rays. Foreign metallic substances in the body are detected by the shadows they cast. The effect of tuberculosis in the lungs can be detected, because dense calcified scars are left on the walls of the lung cavity.

533. X-ray tubes. In Fig. 334, where a well-known type of X-ray tube is shown, *C* is the cathode which focuses the cathode rays on the target, or anticathode, *A*. The X-rays are emitted in all directions from the part of the anticathode that is struck by the cathode rays. B is the anode. The anticathode is mounted on a massive metal support, so that the heat developed by the impact of the rays can be readily dissipated.

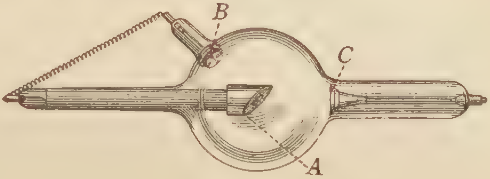


Fig. 334. An X-ray tube

The Coolidge tube is an improved form of X-ray tube. Its principal distinguishing feature is that advantage is taken of the fact that white-hot metals emit electrons in large numbers. The cathode consists of a short strip of tungsten kept white-hot by an electric current, thus causing a copious emission of electrons. This tube has many advantages over the simple one shown in Fig. 334.

534. Nature of X-rays. It is now known that X-rays are *waves*, similar to light-waves but with much shorter wave-lengths.* The fact that light-waves are so extremely short has always been astonishing to the student. The shortest wave-length of light which affects the eye is about 0.00004 centimeter, but this is several thousand times longer than waves of X-rays. It is this extreme shortness of wave-length, with the corresponding high frequency, which gives X-rays such great penetrating power.

X-rays do not all have the same properties; some are not so penetrating as others. Those having greater penetrating power have shorter wave-lengths and higher frequencies. In practice the quality of the X-rays can be changed by changing the speed of the cathode rays which cause the anticathode to radiate these rays. The greater the velocity of the cathode rays, the more penetrating the X-rays they produce. The velocity of cathode rays is changed by changing the voltage applied to the tube. In general, metals of high atomic weights can be made to give X-rays of shorter wave-length or more penetrating rays than those given off by lighter substances.

---

* A method for measuring the wave-length of X-rays is explained in section 659, page 697.

535. Electronic emission used in the rectification of alternating currents. The stream of electrons which constitute the cathode rays is only one case of the transport of free electrons through space; for example, when ultraviolet light, the frequency of which is higher than that of visible light, falls on polished zinc or some other "electropositive" metal, electrons will be shot off from the surface of the metal. That hot metals will also emit a stream of electrons has already been referred to. In recent years there have been developed a number of practical applications of these electronic streams. Among them is a commercial device for the rectification of an alternating current, the *thermionic rectifier*, sold under the name of the *tungar rectifier*. By the term *rectification* is meant the changing of an alternating current (one which flows first one way and then the other) into a current which flows in only one direction. The principal commercial use of rectifiers is in the charging of storage batteries. For reasons that should be clear to the student, a storage battery can be charged only by a current which flows in one direction. Since the electric current supplied by public-service lines is usually alternating, most public garages and service stations where storage batteries are cared for must install some sort of rectifier. Since the principle involved in the thermionic rectifier is coming into wider use, a brief description of it will be given here.

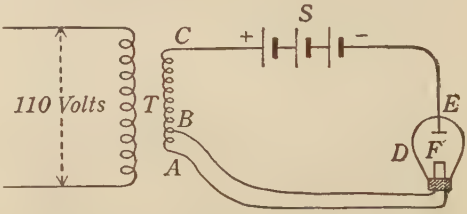


Fig. 335. A thermionic-rectifier circuit

In Fig. 335, *T* is a step-down transformer, the high-tension side of which is connected to the alternating-current lighting circuit, say of 110 volts. The secondary has a wire, *B*, tapped in, so that between *A* and *B* there will be a few volts, enough to heat to incandescence a short tungsten filament, *F*. The voltage between *A* and *C* is not large enough to make the argon gas in the bulb *D* conducting. But when the hot filament *F* emits electrons at a speed sufficient to begin to ionize molecules of this gas by collision, then current can flow through the tube between the terminal *E* and the filament *F*. The important thing to note is that this electronic emission from *F* cannot take place when *F* is connected to the positive and *E* to the negative terminal of the transformer; for then the positive charge on the filament holds back by its attractive force the negatively charged electrons, and the negative charge on *E* repels them. But when *F* is negative and *E* is positive, there will be a stream of electrons flowing out from *F*, which will make the gas a temporary conductor, and a current will flow through the gas from *E* to *F* (this is the direction of flow of the *positive* charge). Hence, as the voltage from the secondary of the transformer changes back and forth, first in one direction and then in the other, current will flow through the tube when *F* is negative and *E* positive; that is, about half the time. The current through the tube will be a pulsating one, but it will be always in the same direction. A storage battery, *S*, in series with the tube, can be charged by this unidirectional pulsating current.

If the tube were exhausted to a very high vacuum, it would still act as a rectifier, but only a relatively small current would flow across—that carried by the electron stream. An inert gas (one that will not attack the hot filament) is used in the tube to supply ions so that a larger current will flow. A medium-sized tungar bulb will carry a charging current of about 6 amperes.

A similar tube,—exhausted, however, to a very high vacuum,—called the *kenotron* (from the Greek *kenos*, meaning "empty"), is used to rectify high-voltage alternating currents.

536. Electronic emission used in amplifiers. Vacuum tubes have been devised that will amplify weak alternating or pulsating currents, such as the current in a telephone circuit. These tubes are usually exhausted to a high vacuum, and contain three elements instead of two (as in the tungar rectifier of Fig. 335). The third terminal is a wire mesh, or "grid," *G* (Fig. 336), which is placed between the hot filament *F* and the cold plate *P*. The plate *P* is positively charged by a battery, *B*. The tube which contains *F*, *G*, and *P* is not shown in the figure. If the grid *G* were not present, the positively charged plate *P* would attract the electrons emitted by the heated filament, and a continuous current—the "plate" current— would flow in the circuit *PBF*. A little consideration will show that the grid may produce the following different effects:

1. If the grid is charged negatively, the charge will repel the electrons from *F* and tend to prevent them from reaching the plate *P*, thus decreasing the current in the circuit comprised of the plate, the battery *B*, and the filament.

2. If the grid is given a small positive charge by some outside agency, the attractive force this charge exerts on the electrons will help the charge on *P* in moving electrons away from the filament, thus increasing the current in the plate circuit.

3. Since the grid is near the filament, a small change in its charge will produce a relatively large change in the plate current.

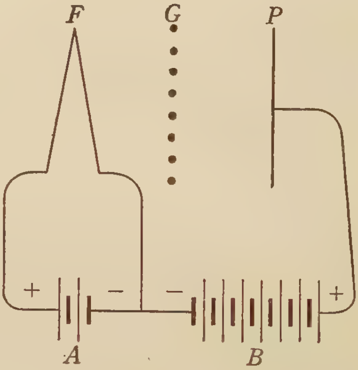


FIG. 336

The current in the plate circuit is thus very sensitive to small changes in the charge (or potential) of the grid. Moreover, it requires only a negligible amount of energy to charge the grid. When the grid is given a positive charge, the current in the plate circuit is increased, but the energy for this increased current is obtained from the battery *B*. This gives the necessary condition for the tube to act as a relay, or amplifier. This will be clearer to the student in the application which follows.

In long-distance telephony, vacuum-tube amplifiers, or relays, are used to increase the energy in the telephonic currents. Fig. 337 shows, in a simplified manner, the connections.*

The weak telephonic current, which has come from a great distance, flows through the primary of the transformer T_1. One end of the secondary of this is connected to the grid, the other to the filament. As in Fig. 336, the battery *A* supplies the current for heating the filament, and the battery *B* is used to supply the current in the plate circuit. The telephone current in the primary of the transformer T_1 causes fluctuations in the potential of the grid, and these, in turn, cause relatively large changes in the current of the plate circuit. The primary of the transformer T_2 is connected in the plate circuit. When the current in this circuit fluctuates, induced currents are produced in the secondary of T_2. This current is the amplified telephone current. It has been found practicable to produce this amplification without any appreciable distortion* of the sound heard at the receiving end. Moreover, it has also been found practicable to deliver the current from the secondary of T_2 to another tube circuit, where it is again amplified.

---

* The connections shown here are obviously incomplete, for this method would permit "talk" to be transmitted in only one direction.

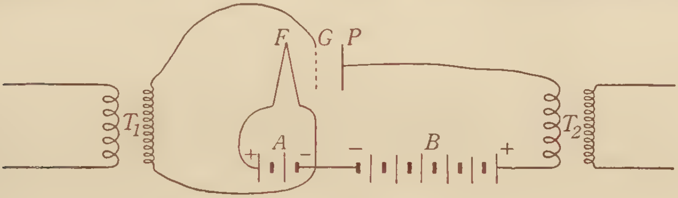


FIG. 337. A thermionic-amplifier circuit

In actual practice a number of amplifiers can be connected in series. The discovery of this vacuum-tube amplifier was one of the important factors in making transcontinental wire telephony possible.

Vacuum-tube amplifiers are used to enable very large audiences to hear a speaker. By using a large number of them the sound may be amplified a million times. It is now possible for a speaker to address a large audience a thousand miles away. He delivers his speech into a transmitter of a long-distance telephone line, and at the other end the electric currents are amplified and delivered to a number of loud-speaking instruments.

537. Emission of electrons. The ionization of air or other gas is so often due to the impact, or collision, of electrons with gas molecules that it is well to review here the principal ways in which free electrons can be produced.

---

*A relative change in the intensity of the overtones of a vowel sound may change this sound to one of a different vowel (see section 327).

1. *Ultraviolet light.* When waves of light shorter than those producing the sensation of violet (*ultraviolet* waves) fall on any substance, they may cause that substance to emit electrons. It is most marked in the case of the electropositive metals, such as zinc, sodium, and potassium. If a piece of polished zinc is connected to an electroscope and charged with negative electricity, the electroscope shows that the charge is rapidly lost when the zinc is exposed to light from a naked carbon arc, which is rich in ultraviolet light. This is due to the emission by the plate of negatively charged electrons.* But if the zinc is charged positively, it will not lose its charge when illuminated. This is because the positive charge on the zinc, attracting negatively charged electrons, prevents them from leaving the zinc. No ions are produced which can carry away a positive charge.

If ultraviolet light falls on the terminals of a spark gap, a spark can be produced by a lower potential than when no light is used. Why? It is necessary to illuminate only the negative terminal.

The effect produced by ultraviolet light is often called the *photoelectric* effect.

2. *X-rays.* The atoms of all substances—whether solid, liquid, or gaseous—emit electrons when they are illuminated by X-rays. The ionization of gases produced by X-rays is an indirect effect. X-rays cause the emission of electrons at high speeds from molecules of the gas, and these electrons ionize the gas by collision with other molecules.

3. *Radiations from radioactive substances.* The radiations from radium and other radioactive substances cause electrons to be emitted from atoms which these radiations strike.

4. *High temperature.* The emission of electrons from hot bodies has been referred to before, in sections 533 and 535. This is an important method and has many useful applications.

538. Millikan's method for measurement of the fundamental charge of electricity. It has been stated before that Millikan showed that small charges are an exact multiple of a certain quantity,—the smallest charge of electricity which apparently exists. This we have called the *natural* unit (sect. 391). According to Millikan no particle or ion can have a charge which is a fraction of this unit: it must be one or two or three or four etc. of these units. The charge on an electron is always one of these units.

---

* Glass is opaque to ultraviolet light; hence a sheet of glass held between the arc and the negatively charged zinc plate will stop the emission of electrons and thus stop the loss of charge.

It is worth while to give a simple explanation of Millikan's method of measuring small charges. A tiny drop from a spray produced by an atomizer was permitted to fall through a small hole into the space between two horizontal plates. Such a droplet usually picks up a charge by colliding with an ion in the air (ions can easily be produced in the air by the radiation from radium). The droplet was watched by means of a low-power microscope, and its rate of fall carefully measured. From the rate at which the drop fell through air under the action of gravity alone, its size could be determined (sect. 110), and, the density of the material of the droplet being known, its mass could be computed. Then the two parallel plates were charged with electricity so as to drive the droplet up or down, as desired. From the velocity produced by the charged plates the electrical force on the droplet could be computed. The electric force acting was the strength of the electric field times the charge on the droplet (see equation (9), sect. 392). Since the strength of the field was known, the charge of the droplet could be calculated. The actual details, as may be expected, were much more complicated than this brief description indicates.

In several thousand experiments Millikan found that *the charge on these droplets was always a whole number of these unit charges* where the unit charge was 4.774×10^{-10} electrostatic units, or 1.591×10^{-19} coulombs. We now believe that all charges, whether large or small, are multiples of this unit.

One result of Millikan's determination of the value of e, the unit charge, is that it gives us (as explained in section 455) an accurate method of determining the masses of atoms. Similarly, if the value of e/m for the electrons in the cathode-ray stream can be determined, the mass of the electrons can be calculated. (The ratio of e/m for an electron can be determined by measuring the deflection of the rays by both a magnetic and an electric field. The theory of this is too involved to be given here.) This is the method by which it was found that the mass of the electron is about one eighteen-hundredth of the mass of the smallest known atom—that of hydrogen.

CHAPTER XXXVIII

RADIOACTIVITY

Introduction, 539. Types of radiation emitted, 540. Effects of electric and magnetic fields, 541. Atomic disintegration, 542. Radioactive substances in the uranium-radium series, 543. Source of the gamma radiation, 544. Structure of an atom, 545. Atomic numbers, 546. Isotopes, 547. Other radioactive series, 548.

539. Introduction. The discovery of X-rays attracted a great deal of attention, not only among scientists but also among the general public. It was apparently an entirely new type of radiation, and hence of high theoretical importance. Moreover, the great practical benefit from the use of this radiation was immediately recognized. Naturally there was much speculation as to just what there was in the vacuum tube that caused this radiation. Since the glass walls of all vacuum tubes which emitted X-rays exhibited visible phosphorescence, it probably occurred to many that the X-ray radiation was in some way associated with this phosphorescence. Becquerel examined minerals and other substances which phosphoresced under the action of light, hoping that they too might give an invisible radiation. In this way he found, in 1896, that uranium and its compounds gave an invisible radiation capable of affecting a photographic plate through black paper or through sheets of aluminum. Becquerel soon found that this new radiation had nothing to do with the phosphorescence: that it could be obtained from nonphosphorescent uranium salts. He found also that these radiations, like X-rays, made air, or any gas which they traversed, electrically conducting.

Soon after Becquerel's discoveries M^{me} Curie and G. C. Schmidt independently found that *thorium* and its compounds also emitted the same sort of radiation as that given out by uranium. A little later M^{me} Curie found in pitchblende some traces of a new element, which she called *radium*. Radium has since been found in a number of ores, one of the best sources at present being the carnotite found in the Rocky Mountains. Radium emits a radiation a million times more intense than that from uranium. Still later, in 1900, Debierne discovered another element, *actinium*, which emits a radiation similar to that from radium.

It was soon discovered that the ionizing of the air by these radiations afforded an extremely delicate test of their presence. If one of these radioactive substances is brought near a charged electroscope, the electroscope is discharged by the conducting air which surrounds it. The rate at which it is discharged gives a measure of the intensity of the radiation. The method can be made so sensitive that feeble radiations, which cannot be detected by any other method, are easily observed.

The radiations from radium are commonly used to produce phosphorescence of zinc sulfide. One thousandth of a milligram of pure radium mixed with zinc-sulfide crystals will be enough for the luminous paint on a watch face.

540. Types of radiation emitted. It has been found that there are three different types of radiation emitted by radioactive substances. These have been called the *alpha* (α), *beta* (β) and *gamma* (γ) radiations.

Alpha rays. The *alpha* radiation has been found to consist of small particles of matter shot out at speeds of about twenty thousand miles per second. The mass of each of these is equal to that of a *helium* atom, or about four times the mass of a hydrogen atom. They are always charged positively with two of the natural units of electricity.

Alpha rays have a very great ionizing effect on air. They cannot, however, penetrate more than a few centimeters of air at ordinary pressures. Their great ionizing power is probably due to their relatively large momentum. Rushing through the air at high speeds, they are able to knock off electrons from a large number of atoms in their path and thus to produce large numbers of charged ions.*

It has been found that radium is a constant source of heat, a gram of radium giving off about 130 gram-calories per hour. This heating effect is due mainly to the alpha particles. The relatively large amount of kinetic energy of these particles is converted into heat when their motion is stopped within the mass of radium itself, or by the inclosing vessel. The source of this energy must have been in the radium atom before the alpha particle was shot out. This gives a hint of the enormous amount of internal energy that exists in atoms.

---

*It should be remembered that when a neutral atom loses an electron, it becomes positively charged.

When an alpha particle strikes a phosphorescing substance, such as zinc sulfide, it produces a tiny flash of light.* It has even been found possible to count the number of alpha particles emitted from an active substance by counting the flashes of light on a zinc-sulfide screen. They can also be detected by their photographic action on a sensitive plate.

Beta rays. Beta rays are found to consist of streams of negatively charged electrons moving at high speeds. The speed is different in the case of different radioactive substances; in the case of radium and its products these speeds range from about three tenths to nine tenths of the velocity of light.

Beta rays are the same as the cathode rays of a vacuum tube, except that in general they have higher speeds. When they strike a substance, they cause it to emit a radiation similar to that of X-rays. They produce phosphorescence and affect a photographic plate, and their penetrating power is much greater than that of the alpha rays.

Gamma radiation. Gamma radiation consists of very short waves, similar to X-rays, but in general of higher frequency and shorter wave-length than X-rays. It is the most penetrating radiation known, and is intense enough for detection after traversing about a foot of iron. This radiation also produces phosphorescence and affects a photographic plate. It causes all objects which it strikes to shoot off electrons. Gamma rays produce relatively only a slight ionizing effect in gases; this seems to be a secondary effect, the ionization probably being produced by the electrons which the gamma radiation causes the gas atoms to emit.

541. Effects of electric and magnetic fields. The effect of an electric field on the radiations from a radioactive substance can be shown by the method illustrated in Fig. 338. *A* and *B* are two parallel plates which, when charged as shown, have an electric field between them. The radioactive substance is placed at the bottom of a round, deep "well" in a block of lead. On account of the screening action of the lead, only a narrow beam of radiation is emitted. Under the influence of the electric field the three different types of radiation in the emitted beam are separated as shown. If a photographic plate is placed at *PP*, each bundle of rays will produce a black spot on the plate, so that the deviation of the alpha and beta rays can be measured. The student should verify the directions of the deflections by recalling that the alpha and beta rays consist of particles carrying positive and negative charges respectively. The gamma radiation, being a wave-motion, should not be affected. The figure does not show the relative sizes of the deflections, for the beta particles are deflected very much more than the more massive alpha particles.

---

* These flashes may be seen by looking at the luminous dial of a watch through a magnifying glass.

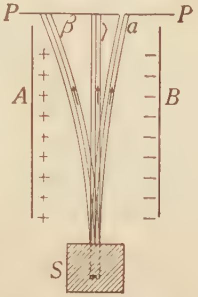


FIG. 338. The deflection of the radiations by an electric field

Fig. 339 shows the deflection produced by a magnetic field which is perpendicular to the plane of the paper and directed toward the student. When the magnetic field is in this direction, the rays are separated as shown. Why?

The measured magnitudes of the deflections of the alpha and beta radiations in electric and magnetic fields of known strengths give the data for computing the masses and the velocities of the particles. Hence a more complete study of these methods is of great interest to advanced students.

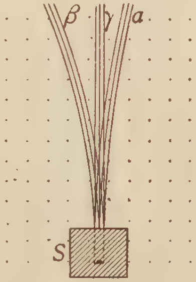


FIG. 339. The deflection of the radiations by a magnetic field

542. Atomic disintegration. The student should recall that the atom of any element is believed to consist of a small positively charged nucleus surrounded by negatively charged electrons. Different kinds of atoms differ from one another in the mass and construction of the nucleus and in the number of the surrounding electrons. The understanding of the details of this structure is one of the great problems of the day. However, even simple considerations will show that in the atom some internal arrangements will be far more stable than others. In none of the more common elements is there any evidence of changes in the structure of these atoms; hence they must be in some very stable condition. On the other hand, the atoms of the radioactive elements are in a relatively unstable condition. They have a tendency to rearrange their internal structure and, while they are doing this, to throw off an electron or some other part of the nucleus. Apparently a very large amount of potential energy is stored in each of these atoms; and when the change in structure takes place, the change in the internal energy is large enough to cause the rearrangement to take place with explosive violence. The parts shot out thus carry with them considerable energy.

Imagine a specimen of *pure* radium; that is, one in which all the atoms are of radium. All of them are unstable; but for any one atom it is a matter of chance when the explosion takes place. It will take place in one atom here, in another there, and so on. Since there are so many million atoms in our sample, the rate of these explosions is nearly a constant. It is something similar to the death rate of a large city: although in small groups deaths occur very irregularly, yet for an entire city the rate is fairly uniform. Estimates made for radium indicate that, starting with a quantity of pure radium, half the atoms will be disintegrated in about sixteen hundred years, half the remainder in another sixteen hundred years, and so on, this time interval being called the *period* of radium. When an atom of radium explodes, it shoots off an alpha particle from the nucleus and becomes an atom of a different element. The new element is called *radium emanation.* When an atom of radium emanation explodes, it emits another alpha particle and becomes the atom of an element called *radium A*; this, disintegrating in turn, emits an alpha particle and becomes *radium B*; when the atom of radium B disintegrates, it emits a *beta* particle and becomes *radium C*; and so on. The complete story can be traced out from the table on page 576. It is thus seen that the sample of pure radium with which we started has in a short time become a mixture of different kinds of atoms: radium, radium emanation, radium A, etc.

One of the most interesting questions today about radium and its products is, What is the *end* product of radium,—the last element in this series? All evidence points toward *lead* as the end product; that is, a sample of pure radium will in time all be converted into lead, a stable element. Such a change would require billions of years. While such changes cannot be directly observed, there is much evidence in favor of this belief; for example, the fact that lead is found in all ores containing radium.

Radium itself is believed to be a descendant of uranium. The evidence for this is quite definite, although uranium disintegrates so slowly that direct proof is difficult. Perhaps the first definite indication came from the fact that radium is found in all ores containing uranium, and the amount of radium in any given ore is always proportional to the amount of uranium in that ore.

543. Radioactive substances in the uranium-radium series. The table on page 576 gives a fairly complete story of the radioactive materials in the uranium-radium series. It begins with uranium, the heaviest known atom. When an atom of uranium disintegrates, it shoots off an alpha particle and becomes an atom of another element, called *uranium* X_1; when this disintegrates in turn, it emits a beta particle and becomes an atom of *uranium* X_2. Thus, reading down the first column of the table, one finds the names of the elements in the order of their transformation. Each is the descendant of the one preceding and the parent of the one following.

An atom of an element radiates only once while it is an atom of that element. Each radiation is the result of an explosion, after which it at once becomes an atom of another element—the next in the series. Then at some later time (a time apparently dependent on chance) it again disintegrates and becomes an atom of the next element, and so on.

The second column gives the "period" of the element,—the estimated time required for half of a very large number of the atoms to become disintegrated. Obviously a long period means a slow rate of disintegration, and a short period a rapid rate.

THE URANIUM-RADIUM SERIES

ELEMENT	PERIOD	RADIATION	ATOMIC MASS (THEORETICAL)	ATOMIC NUMBER
Uranium	5×10^9 yr.	α	238	92
Uranium X_1	23.8 da.	β and γ	234	90
Uranium X_2	1.15 min.	β and γ	234	91
Uranium II	2×10^6 yr.	α	234	92
Ionium	9×10^4	α	230	90
Radium	1580 yr.	α	226	88
Radium emanation	3.85 da.	α	222	86
Radium A	3.0 min.	α	218	84
Radium B	26.7 min.	β and γ	214	82
Radium C	19.5 min.	β and γ	214	83
Radium C_1	Very short	α	214	34
Radium D	16.5 yr.	β and γ	210	82
Radium E	4.85 da.	β and γ	210	83
Polonium	136 da.	α	210	84
Lead	——	——	206	82

The atomic mass of radium is about 226. When an atom of it explodes, an alpha particle with a mass of 4 is shot out. The atom of the element called radium emanation, which results from this explosion, should then have an atomic mass of about 222. Thus, whenever an atom in disintegrating loses an alpha particle, its atomic mass decreases by 4. If it loses a beta particle, there should be no appreciable change, for the mass of a beta particle is too small to produce any measurable effect. On the whole, experimental evidence supports these explanations.

In the transition from uranium to lead there are eight transformations in which an alpha particle is shot out; hence there should be a total loss in atomic mass of $8 \times 4 = 32$. The accepted value of the atomic mass of uranium is 238. Subtracting 32 from this we get 206, the atomic mass of lead which has descended from uranium. The atomic mass of ordinary lead is 207.2. It is now known that there is more than one kind of lead, and that ordinary lead is a mixture. (The different kinds of lead have identical chemical properties, but different atomic masses.) Lead found in radioactive ores has an atomic mass of 206.3. This agrees well with the theoretical value, 206, given above.

544. Source of the gamma radiation. When a stream of cathode rays strikes any substance, X-rays are emitted by that substance. In a similar manner, when a stream of beta rays (electrons of higher speeds than those given off from a cathode) strikes any substance, it causes the substance to emit a gamma radiation, which is a very penetrating form of X-rays. Since the beta radiations from a radioactive substance originate not only at the surface but all through the interior of the mass, a large percentage of the beta particles ejected from the atoms strike adjacent atoms of the radioactive substance, producing oscillations in these atoms which cause them to emit the wave-motion called gamma rays. High-frequency waves should also be caused by the emission of a beta particle from the nucleus of an atom; for undoubtedly violent oscillations within the atom must be produced. Hence we should expect gamma radiation in all those transformations where beta rays are emitted. This is found to be the case.

545. Structure of an atom. It has been stated several times in this book that an atom is believed to consist of a positively charged nucleus with negatively charged electrons clustered around it. This conclusion has been reached largely through a study of the radioactive properties of substances and their radiations.*

The fact that the alpha and beta particles are shot out with so much energy indicates that they originate in the nucleus. Since the alpha particles from all radioactive substances have the same mass (that of the helium atom), and the beta particles are always electrons, it follows that the central part of the atom is probably made up of an assemblage which contains positively charged alpha particles and negatively charged electrons.† But there must be an excess of positive charge in the nucleus. Surrounding the nucleus are electrons. The number of these is determined by the magnitude of the excess of positive charge on the nucleus. In an uncharged atom there must be just enough electrons to neutralize the positive charge on the nucleus; for example, if the excess positive charge is 20 natural units, then there must be normally 20 electrons, since each carries 1 natural unit of negative electricity. It has thus been possible to estimate the number of the surrounding electrons. It is generally believed that the hydrogen atom has only 1, and that the number for other elements is about half their atomic mass and equal to the "atomic number," which is explained in the next section.

---

* Confirmation is found in the study of the spectra of light and X-rays.

† To account for the fact that the masses of the different kinds of atoms are not always an integral multiple of the mass of the alpha particle, it is usually assumed that the nuclei of some elements contain also hydrogen nuclei. There is some evidence in favor of this assumption. Rutherford recently has been apparently able, by using swift alpha particles, to knock hydrogen nuclei out of the nuclei of a number of different kinds of atoms. All these atoms have a mass which is not an integral multiple of the mass of an alpha particle.

Certain experiments indicate that the diameter of an atom is about 10^{-8} centimeters and that the diameter of an electron is about 10^{-13} centimeters. The diameter of the nucleus of a gold atom, which is relatively massive, is not over 10^{-12} centimeters. The diameter of the nucleus of a hydrogen atom is estimated to be about 10^{-16} centimeters.

546. Atomic numbers. As is well known, there is a certain progressive change in the properties of chemical elements. In texts on chemistry one will usually find a table with the elements arranged in accordance with the "periodic law." If these elements are numbered consecutively hydrogen 1, helium 2, lithium 3, etc., numbers will be obtained which are called the *atomic numbers*. In most cases these numbers follow the order of the atomic masses.

The character of X-ray radiation given off from any element struck by a stream of high-speed electrons has been found to depend on the atomic number of that element; in fact, a study of the characteristics of the X-rays emitted by different elements gives the most accurate method of determining the atomic number of many elements.

According to current theories of the structure of the atom, the atomic number of any element is always equal to the number of natural units of the excess of positive electricity on the nucleus of the atom of that element. This is also equal to the number of electrons surrounding the nucleus when the atom is in electrical equilibrium, because in that case there must be a sufficient number of electrons to neutralize the positive charge on the nucleus.

547. Isotopes. The chemical properties of an element depend on the atomic number of that element, and this apparently depends on the electrical charge on the nucleus. Since the charges carried off by the radiation in the disintegration of radioactive substances are known, it is possible to compute the changes in the charge on the nucleus. These changes are traced out in the following table for a portion of the uranium-radium series.

ELEMENT	NET CHARGE ON NUCLEUS	RADIATION	CHARGE LOST	RESULTANT NET CHARGE	RESULTANT ATOM
Radium A	+ 84	α	+ 2	+ 82	Radium B
Radium B	+ 82	β	− 1	+ 83	Radium C
Radium C	+ 83	β	− 1	+ 84	Radium C_1
Radium C_1	+ 84	α	+ 2	+ 82	Radium D
Radium D	+ 82	β	− 1	+ 83	Radium E
Radium E	+ 83	β	− 1	+ 84	Radium F
Radium F	+ 84	α	+ 2	+ 82	Lead
Lead	+ 82	——	——	——	——

The assumption is made, in compiling this table, that the charge on the nucleus of radium A is equal to the atomic number 84. With this assumption one can, by the use of the table, readily arrive at the other nuclear charges, since the charge carried away from an atom by the radiation is known. Two things become evident from the table:

1. There is a confirmation of the theory that the atomic number is equal to the nuclear charge, for the computed nuclear charges in this table agree with the values of the atomic numbers which are known.

2. Certain of these elements have the same nuclear charge and the same atomic numbers. Radium B, radium D, and lead have an atomic number 82. They are found to have *identical chemical properties*. Radium A, C_1, and F have an atomic number 84. They too have the same chemical properties.

Elements having the same chemical properties are called *isotopes*.

It has been recently found by Aston that there are a large number of isotopes. Thus chlorine is really a mixture of two gases: one having an atomic mass of 35.0, the other having an atomic mass of 37.0. These two gases have identical chemical properties. It is now believed that many of the so-called elements are really mixtures of isotopes.

548. Other radioactive series. The uranium-radium-lead series is not the only known series of elements that are continuously disintegrating. Actinium, which is probably another descendant of uranium, gives rise to a number of other elements. Thorium and its descendants apparently form a series independent of that of uranium. They have about the same range in atomic masses and atomic numbers as those of the uranium-radium series, and have as the end product lead with an atomic mass of 208,—an isotope of the lead which is the end product of the uranium-radium series. Both the actinium and the thorium series give out the same kind of radiations as those of the uranium-radium series.

CHAPTER XXXIX

HIGH-FREQUENCY OSCILLATIONS AND ELECTRICAL WAVES

A mechanical oscillating system, 549. An electrical oscillating system, 550. A method of producing a series of damped oscillations, 551. Electrical resonance, 552. Undamped high-frequency oscillations, 553. Electromagnetic waves along wires, 554. Electromagnetic waves in space, 555. The detection of electromagnetic waves, 556.

549. A mechanical oscillating system. Let us first consider the case of something which we know will oscillate when disturbed from its position of rest. As a simple case take a mass suspended by a spiral spring as shown in Fig. 340. If the mass is raised or lowered and then released, it will move up and down with an oscillating motion. A little thought will show us that there are three necessary conditions which must be fulfilled in order that the mass may oscillate.

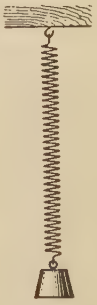


FIG. 340

1. There must be a *return force*; that is, one which tends to restore the mass to its original position. In this case the force is furnished by the spring and by gravity.

2. The moving parts must have *inertia*; for otherwise the mass would not move past its position of equilibrium but, when released, would move immediately to this position and stop.

3. The *friction* must not be too large. If the mass were hung in a viscous fluid, it might not oscillate, but might slowly return to its position of equilibrium.

These three conditions hold not only for the foregoing case but *for all types of oscillation*. If, for any kind of contrivance, these three conditions are met, then we may be certain that there will be oscillations whenever the equilibrium is disturbed.

550. An electrical oscillating system. In view of what was stated in the last section, it is necessary to show that three conditions must be met if a circuit is to have electrical oscillations. Fig. 341 shows a circuit consisting of a *condenser*, C, and a coil of wire, L, having *self-induction*. Suppose that the condenser is charged as shown, and that the circuit is open at the key, K. When the key is closed, the condenser will begin to discharge, producing a current through the coil L. The *self-induction* of this coil gives to this current *inertia*. For when the current begins to flow, there will be an induced electromotive force of self-induction (sect. 487) which opposes the flow; then, after the current gets started, the electromotive force of self-induction opposes the stopping of the current. For the same reason that a mass will swing past its position of equilibrium, the inertia (self-induction) of the current prevents it from stopping at the exact moment the condenser is fully discharged. The current, continuing, charges the condenser in the opposite direction; then it discharges in the first direction; and so on.

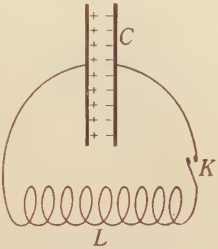


FIG. 341

The *condenser* in the circuit supplies the "return force" of the first condition, and *self-induction* the inertia of the second condition. The "friction" is determined by the electrical *resistance* of the coil and connecting wires. If the resistance is large, the energy is dissipated quickly in the form of heat; if the resistance is not too great, the condenser discharge is an oscillatory one.

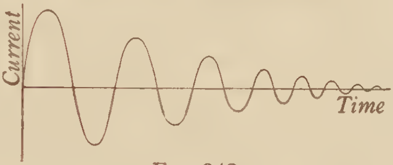


FIG. 342

The curve in Fig. 342 represents an oscillatory current in the coil L. Distances above the horizontal line mean current in one direction; those below the line, current in the opposite direction. The "damping," or diminution, in the amplitude of the current is due to the resistance of the circuit. An oscillation of this type is called a *damped oscillation.*

For the case where the resistance is relatively small it is shown in more advanced texts that the period of vibration T is given by the equation

$$T = 2\pi\sqrt{LC}, \tag{103}$$

and the frequency f by the equation

$$f = \frac{1}{2\pi\sqrt{LC}}, \tag{104}$$

where L is the inductance of the coil, and C is the capacitance of the condenser.

In the case where a telephone condenser of 1 microfarad capacitance and a single-layer coil of wire consisting of 100 turns wound closely on a cylindrical surface 20 centimeters in diameter ($L = 0.0023$ henry) are used, the period of oscillation is about 0.0003 second, and the frequency about 3300 per second.

A circuit containing capacitance, inductance, and low resistance always has a natural period of oscillation, just as definite as the period of a tuning-fork or of a simple pendulum. Practically all circuits have some inductance; indeed, it would be very difficult to design a circuit which would not have some magnetic flux interlinking it when a current flowed through it. Many circuits have capacitance, although at first this is not obvious. For example, wireless aërials, now so commonly seen, have capacitance. In this case the aërial forms one "plate" of the condenser, and the ground the other. When the aërial is charged positively, there will be an induced negative charge on the ground; when it is negatively charged, the induced charge is positive. If such an aërial is charged and then connected to the ground, it will discharge with an oscillatory current. The natural frequency of the aërials used by amateur wireless operators is about 1,500,000 vibrations per second. The frequency is high because both the capacitance and the inductance are small.

551. A method of producing a series of damped oscillations. When a condenser is charged and then discharged with an oscillatory current, the oscillations quickly die out. Often it is desired to have a rapid succession of such discharges. The principle of the method used is simple. In Fig. 343 a diagram of the connections of the apparatus is given. A transformer, T, is connected on the low-tension side to the ordinary alternating-current supply; the high-tension side, giving at least 10,000 volts, is connected to the condenser *C*. A spark gap, *S*, is inserted in the discharge circuit. When the condenser gets charged to a sufficiently high voltage, a spark will jump across this gap, making it a conductor for a sufficient time to enable the oscillatory discharge to pass. When the oscillatory current dies out, the spark gap automatically becomes a nonconductor, and the condenser can be charged again. If the primary of the transformer is connected to a 60-cycle supply, the condenser will be fully charged 120 times per second, and 120 sets of oscillatory discharges will be produced each second (if there were no spark gap in the circuit, the condenser would never become highly charged, as it would be short-circuited by the coil *L*).

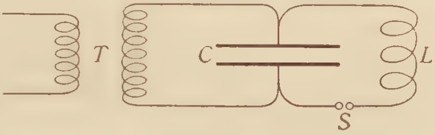


FIG. 343

Fig. 344 shows the kind of oscillations produced in such an apparatus. There is a discharge of damped oscillations; then, after a short interval, another oscillatory discharge; then another; and so on—often several hundred each second. Each group in the figure consists of the high-frequency oscillations, the frequency of which depends on the capacitance and the inductance of the circuit. The number of groups each second is equal to the number of spark discharges each second at the gap.

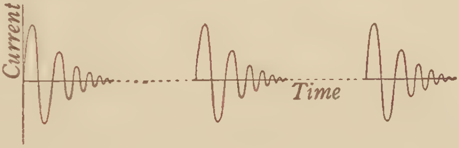


FIG. 344

Those familiar with such apparatus know that special contrivances, such as rotary spark gaps, are often used to give a definite number of discharges each second and to prevent troubles which are caused by the overheating of the terminals of the gap. The method given here is essentially that used in small wireless stations for producing high-frequency damped electrical oscillations.

552. Electrical resonance. The term *resonance* is usually applied to any case where vibrations are produced by an agent which has the same frequency as the vibrator. There are many well-known cases in musical instruments; for example, a string or a tuning-fork can be set in vibration by a sound-wave which has the same frequency of vibration. Cases of resonance are by no means confined to sound apparatus; there are many other common cases. If two simple pendulums of the same length are hung from the same support, and one is set in vibration, the second one will be set in motion by resonance if the support is not too rigid. An interesting case, now used a great deal, is that of electrical resonance.

Fig. 345 shows an arrangement of apparatus by which electrical resonance can be shown very clearly. Oscillating discharges are produced in the circuit consisting of the condenser C_1, the spark gap S, and the coil L_1. The other circuit, containing the condenser C_2, the coil L_2, and an ammeter, A, of the hot-wire type, is not connected to the first circuit, but the two coils L_1 and L_2 are near enough so that part of the magnetic field of L_1 interlinks L_2. The coils $L_1$and L_2 should have a few turns of a diameter of 20 or 30 centimeters. Each of the condensers C_1 and C_2 may be a pair of Leyden jars connected in parallel.

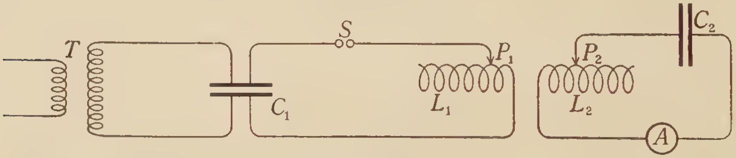


FIG. 345

The oscillatory discharge through the coil L_1 will produce an oscillating magnetic field which induces in the coil L_2 an electromotive force of the same frequency as the oscillations in the C_1L_1 circuit. When the two circuits are adjusted to the same frequency by changing the contact points P_1 and P_2, a relatively large current will be induced, as shown by the ammeter A. If all connecting wires are large, so that the resistance in each circuit is small, and if the coils L_1 and L_2 are not too close together, the "tuning" will be quite sharp; that is, the current in A will be a maximum for one position of the contact P_2 and will be decreased considerably by small changes in the position of P_2.

The principle of the experiment may be stated as follows: The current in the circuit C_1L_1 oscillates with a definite frequency. When the contact P_2 is moved, the inductance of the coil L_2 is changed, thus changing the natural frequency of the circuit C_2AL_2. When the frequency of this circuit is the same as that of the circuit CL_1, a maximum current is shown by the ammeter A.

553. Undamped high-frequency oscillations. There are a number of ways of producing *continuous* oscillations; that is, those of an *undamped* type. In each of these cases the method is to impress on a circuit containing capacitance and inductance a continuously oscillating electromotive force of the same frequency as the natural frequency of the circuit.

The three principal methods for producing the necessary high-frequency electromotive forces will be given briefly.

The alternating-current generator. The usual alternating-current generator produces rather low frequencies (up to 100 per second), but Goldschmidt and Alexanderson have designed generators which will give high frequencies (up to 200,000 per second). High-frequency Alexanderson generators of large power (200 kilowatts) are now used for producing the electrical oscillations used in long-range wireless telegraphy.

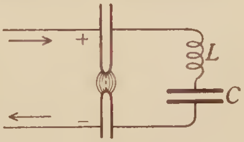


FIG. 346

The arc. If a coil having inductance and a condenser, C, is connected in parallel with a direct-current carbon arc, *sustained* oscillations will, under favorable conditions, be set up in the circuit consisting of arc, condenser, and coil L. If another coil is brought near L, an oscillating electromotive force will be induced in this coil. The cause of these oscillations will not be explained here; it is based on the fact that an increase of current through the arc lowers the difference of potential between its terminals, and a decrease in the current causes an increase of potential—just the opposite of what would happen if an arc obeyed Ohm's law.

The vacuum-tube generator. In section 536 it was shown that slight changes in the potential of the grid of a three-electrode vacuum tube would cause relatively large changes in the plate current. This action can be used to produce continuous oscillations of high frequency. There are many different ways of doing this. In the circuits shown in Fig. 347 the oscillations are produced in the circuit containing the coils M_2 and N_2 and the condenser C. Let us suppose that in some way the plate current, which flows through the tube between the plate P and the filament F and through the battery B and the coil N_1, is increased a little. The increasing current in N_1 will induce in N_2 a current which charges the condenser C, and which, when it flows through M_2, induces in the coil M_1 an electromotive force which changes the potential of the grid. If the proper terminal of M_1 is connected to the grid, the change in the potential of the grid will cause the plate current to increase still more, thus producing a greater effect in N_2 and giving the condenser a larger charge. When the condenser starts to discharge, the induced electromotive force in M_1 affects the potential of the grid, thus changing the plate current in such a way as to induce an electromotive force in N_2 which assists the discharge current. To state it a little differently: As the condenser current oscillates it produces in M_1 an induced electromotive force which affects the potential of the grid, first in one way and then in the other, but always in such a direction that the plate current in N_1 is changed so that it will induce electromotive forces in N_2 in such a direction as to cause the oscillating current in M_2N_2C to get larger and larger, until the rate at which energy is dissipated is equal to the rate at which energy is supplied to it.

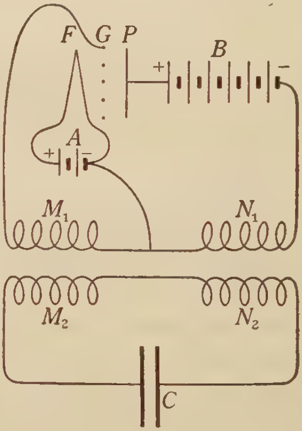


Fig. 347

One very great advantage of the vacuum-tube method is that oscillations having a frequency ranging from a few vibrations to over 1,000,000 per second can be produced. The frequencies are changed by changing either the amount of inductance in the coils or the capacitance of the condenser.

Electrical oscillations produced by this method are now extensively employed for experimental purposes. The best-known use is in wireless telegraphy and telephony.

554. Electromagnetic waves along wires. If a battery is suddenly connected to a long transmission line of two wires by closing the keys KK, a current will flow into the line. The two wires of the line form a condenser which is charged by this current. The charging of the line is not an instantaneous process, but travels as a sort of pulse, or wave, along the line. In Fig. 348 the pulse is shown to have reached MN. To the left of MN the wires are charged as shown, with an electric field extending from one wire to another. Between MN and the battery there is a charging current in the line. This current produces a magnetic field surrounding the wires, which is always perpendicular to the electric field extending between the wires. In the space between the wires this magnetic field is perpendicular to the paper and directed away from the student (use the right-hand rule). Hence in this pulse there are both electric and magnetic fields, and these are at right angles to each other. Both these fields have energy; hence there is a flow of energy as the pulse travels out from the battery. The fact that there are two fields here is only an example of a general rule: *In an electrical wave, or pulse, there are always both magnetic and electric fields at right angles to each other.*

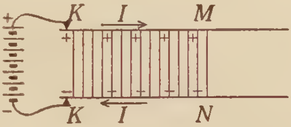


FIG. 348

It can be shown for the simple case where the resistance of the battery and the line is neglected that there are two equations applying to this pulse:

$$I = CEv \qquad (105)$$

and

$$E = LIv, \qquad (106)$$

where I is the current in the pulse, E the electromotive force of the battery (and hence the difference of potential between the two wires), C the capacitance per unit length of the line, L the inductance per unit length of the line, and v the velocity of the pulse.

Eliminating E from (105) by substituting its value as given by (106),

$$I = C(LIv)v = CLIv^2,$$

or

$$1 = CLv^2.$$

Therefore

$$v = \frac{1}{\sqrt{CL}}, \tag{107}$$

an expression giving the velocity of the wave, or pulse. For the case of ordinary wires such as are used in commercial overhead lines, the values of C and L are such that the velocity v is nearly as large as the velocity of light (300,000,000 meters per second).

The pulse of Fig. 348 will travel to the end of the line and there be reflected back. When a guitar string is plucked, pulses travel out from the place where the force was applied to the ends of the string, and, being reflected, run rapidly back and forth. We usually say that the string is vibrating, but the to-and-fro vibrations are caused by the pulses running back and forth. In a similar manner the electrical pulses traveling back and forth on a long transmission line set the line into electrical vibration. In the same sense that a guitar string has a natural frequency of oscillation, so a transmission line has a natural frequency. In section 332 it was shown that strings have different modes in which they can vibrate. In a somewhat similar way a transmission line may oscillate in segments with electrical nodes and loops. The detailed study of the electrical oscillations of lines and the propagation of waves along them is important for the advanced student in electricity, for many interesting things happen in such lines. Telephoning along wires is really a case of the propagation of electrical waves.

While electrical pulses and waves travel along circuits with a definite velocity, a steady electric current does not have a definite velocity. In an electric current the electrons are moving, but the rate at which they move depends on the number participating in the migration and on other rather complex factors.

555. Electromagnetic waves in space. It is now a well-known fact that electromagnetic waves, or, as they are often called, electrical waves, are used in wireless telegraphy and telephony.

In the case of sound-waves there is always a source which is vibrating. This vibrating source causes the air particles to vibrate, and there is propagated a wave-motion consisting of the oscillations of the air particles. In the electrical case the source is some sort of high-frequency oscillation; for example, those of sections 551 and 553. Such a source produces an oscillating electric and magnetic field which radiates as a wave-motion. This may be understood by considering the earliest type of electrical radiator, that of Hertz.

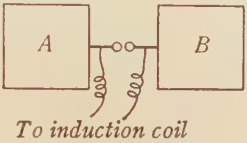


FIG. 349

Hertz's oscillator, or radiator, consists of two sheets of zinc, *A* and *B*, connected by means of a spark gap, as shown in Fig. 349. These plates can be oppositely charged by an induction coil. Although they are not side by side, they serve as the plates of a condenser. When charged to a sufficient difference of potential, a spark will jump across the gap, and the plates will discharge with a damped oscillatory current.

Fig. 350 shows the direction of the *electric field* on one side before the plates discharge. During the oscillating discharge the plates are charged first one way, then the other; hence the electric field in front of the plates changes in direction, oscillating to and fro.

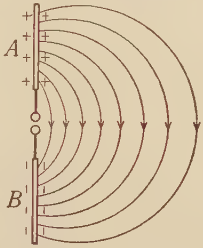


FIG. 350

The current flowing from one plate to the other across the spark gap will produce a *magnetic field* around the spark gap and the connecting rods. Since the direction of the current reverses during the oscillating discharge, this magnetic field will be first in one direction, then in the opposite direction; that is, the magnetic field will also oscillate. The magnetic field will be at right angles to the electric field, as is easily seen by the aid of Fig. 350 and the right-hand rule.

The two oscillating fields (the magnetic and the electric) produced in front of the oscillator move outward, generating what we call an electromagnetic wave. This electromagnetic wave consists of moving electric and magnetic fields, the two fields being at right angles to each other. As a train of such waves passes a given point, the magnetic field is first in one direction and then in the opposite direction, the electric field reversing simultaneously with the magnetic field.

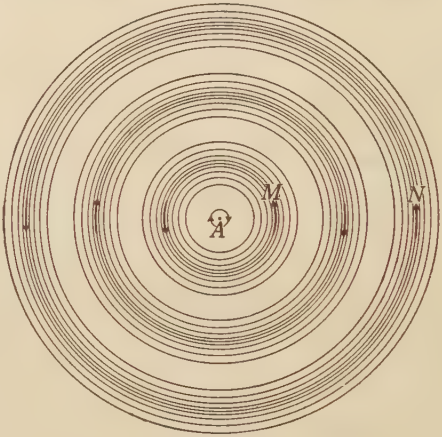


Fig. 351

In the case of a vertical aërial, or wire, in which there is a high-frequency current, the magnetic lines will be horizontal and the electric lines vertical. In Fig. 351 is shown a horizontal section, *A*, of such a wire and its magnetic field. The circles show the position of the magnetic lines at one instant, and the arrowheads give the direction of these lines. The electric lines are vertical, and thus are perpendicular to the magnetic field. A vertical section showing the electric field is given in Fig. 352. These fields are moving out, away from the aërial, with a velocity of 300,000,000 meters per second. The distance from *M* to *N* in each figure (that is, the distance between two corresponding field conditions) is 1 wave-length.

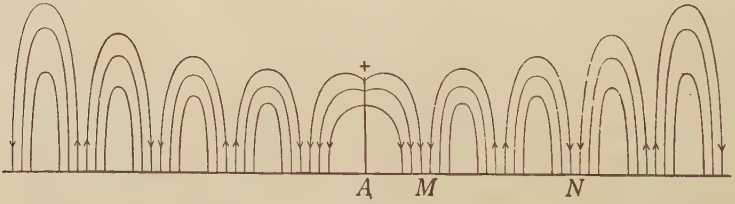


Fig. 352

556. The detection of electromagnetic waves. The details of the various methods for receiving and detecting electromagnetic waves are beyond the limitations of our space, but all the methods depend on two fundamental principles which will be briefly explained.

When an electromagnetic wave, consisting of moving electric and magnetic lines, strikes a conductor, induced electromotive forces are produced. For example, imagine, in the case of Figs. 351 and 352, that there is another vertical wire at some distance from *A*. When the magnetic and electric lines radiated by *A* strike this second aërial, they will induce an electromotive force in it. Further, since the magnetic and electric lines oscillate, having first one direction and then the opposite, the induced electromotive force in the second aërial will act up and then down, oscillating with the same frequency as the current in the first aërial. Thus the first principle involved in the detection of electromagnetic waves is that these waves induce oscillatory electromotive forces in conductors which they strike.

The second principle used is that of resonance. The receiving circuit is adjusted until its natural frequency is the same as the frequency of the incoming wave. When in "tune" the induced currents will be relatively large, and can be measured, or detected, by suitable instruments.

PROBLEMS

1. Compute the natural frequency of oscillation of a circuit consisting of a 0.015-microfarad condenser in series with an inductance of 0.0002 henry.

2. A circuit consisting of a 1-microfarad condenser in series with an inductance has an oscillation frequency of 1000 per second. What is the value of the inductance?

3. A condenser is in series with an inductance of 1.0 henry. What must be the value of the capacitance in order that this combination may be in resonance with the commercial 60-cycle alternating current?

4. What must be the frequency of oscillation of a circuit that produces electromagnetic waves 1000 m. long?

5. What must be the natural frequency of oscillation of a receiving circuit in order that it may be in tune with a 400-meter wave?

PART V. LIGHT

CHAPTER XL

PROPAGATION AND PHOTOMETRY

Radiation of light, 557. The pinhole camera, 558. The inverse-square law, 559. Derivation of the inverse-square law, 560. Photometry, 561. The photometer, 562. Theories of light, 563. Velocity of light, 564. Römer's method, 565. Bradley and the aberration of light, 566. Foucault's method, 567.

557. Radiation of light. Early childhood experience teaches us certain things about light. Among these may be noted the following:

1. Only when a *source* of light is present can other objects be seen. There is a difference between *self-luminous* bodies and those which give light by *reflection*. Most objects are seen by the light which they reflect to the eye.

2. Some objects, like window glass, are *transparent* and freely transmit light; while others—milk glass, for example—are *translucent* (that is, they transmit light, but not so that one can see objects distinctly through them). But the great majority of substances are *opaque*; that is, they do not transmit light except when very thin layers are used.*

3. Light is propagated in sensibly straight lines; that is, light, except in special cases, travels in a straight line from an object to the eye. Measurements in surveying, navigation, and astronomy prove that this is accurately true. This property of light is usually referred to as "the rectilinear propagation of light."

558. The pinhole camera. When light passes through a small opening in a darkened room, inverted images of objects out of doors will be formed on the wall opposite the opening. Light from a candle flame passing through a small hole in a screen will form an inverted image on a second screen (Fig. 353). A box with a small hole in one side is often called a *pinhole camera*. The formation of an image by the light which passes through a small hole is due to the fact that light from different parts of an object travels in practically straight lines through the hole. The student should supply the details and become certain that he understands why an *image* is formed.

---

* Even dense metals transmit light when rolled or beaten into extremely thin layers. One can readily see bright objects through gold leaf.

The shadows of trees cast by sunlight usually have bright patches in them caused by the sunlight coming through small openings between the leaves. The small patches are round *images* of the sun. Why? At the time of a partial eclipse, when the sun is nearly covered, these images are crescent-shaped; that is, of the same shape as the visible part of the sun.

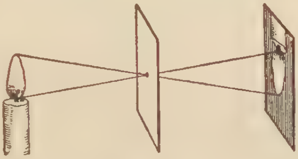


FIG. 353

559. The inverse-square law. The *intensity of illumination* of a given surface is equal to the amount of light that falls on a unit area of the surface in a unit of time. This statement by itself is not clear, for it does not explain what is meant by the expression "amount of light." Sometimes this expression means the amount of energy, measured in ergs, in the radiation; but usually the amount of light is measured in special units which take account of the corresponding sensation produced in the eye. It is not necessary, however, to understand this point fully now. The important part of this definition to be noted now is that intensity of illumination refers to *unit* area.

A very simple law connecting the intensity of illumination of any surface (a book, for example) with the distance of the surface from the source is as follows: *The intensity of illumination of any surface varies inversely as the square of the distance of the surface from the source.*

For example, if a book is held 6 feet from a source of light, and the distance is then reduced to 3 feet, the following proportion is true:

$$\frac{\text{Intensity at 6 feet}}{\text{Intensity at 3 feet}} = \frac{3^2}{6^2} = \frac{9}{36} = \frac{1}{4}.$$

Hence the intensity of illumination at a distance of 3 feet from the source is four times that at 6 feet.

This law is accurately true only when the dimensions of the source are small compared with the other distances involved. For practical purposes the law is sufficiently accurate when the distance of the surface from the source is ten times the largest dimension of the source; for example, if the bright filaments of an incandescent lamp are 2 inches high, the law may be used for distances equal to or greater than 20 inches.

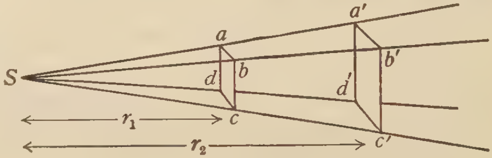


FIG. 354

560. Derivation of the inverse-square law. In Fig. 354 the source of light is at S. Consider only the light which passes through the imaginary rectangular surface $abcd$. Since light travels in straight lines, the same light will pass through the larger area $a'b'c'd'$, where the points a', b', c', and d' lie on straight lines passing through a, b, c, and d respectively and through the source S (for simplicity the surface $a'b'c'd'$ is taken parallel to the surface $abcd$).

Let r_1 be the shortest distance of $abcd$ from S and let r_2 be the distance of $a'b'c'd'$. From the geometry of similar triangles,

$$\frac{ab}{a'b'} = \frac{r_1}{r_2}$$

and

$$\frac{bc}{b'c'} = \frac{r_1}{r_2}.$$

Multiplying these two equations together, we have

$$\frac{ab \times bc}{a'b' \times b'c'} = \frac{r_1^2}{r_2^2}.$$

But the numerator of the left-hand side is the area of the rectangular surface $abcd$, and the denominator the area $a'b'c'd'$; hence,

$$\frac{\text{Area } abcd}{\text{Area } a'b'c'd'} = \frac{r_1^2}{r_2^2}. \tag{1}$$

Let Q be the quantity of light passing through the area $abcd$ and the area $a'b'c'd'$ in 1 second. From the definition of intensity of illumination,

$$\text{Intensity of illumination of } abcd = \frac{Q}{\text{area of } abcd} \tag{2}$$

and $$\text{Intensity of illumination of } a'b'c'd' = \frac{Q}{\text{area of } a'b'c'd'}. \tag{3}$$

Dividing equation (3) by equation (2), we have

$$\frac{\text{Intensity at } a'b'c'd'}{\text{Intensity at } abcd} = \frac{\text{area } abcd}{\text{area } a'b'c'd'}.$$

But the right-hand side is, from equation (1), equal to the ratio of the squares of the distances from the source;

hence $$\frac{\text{Intensity at } a'b'c'd'}{\text{Intensity at } abcd} = \frac{r_1^2}{r_2^2}. \tag{4}$$

This is the *inverse-square law.*

561. Photometry. That branch of the subject of light which deals with the measurement of the luminous intensity of different sources of light and with the illumination produced by various sources is called *photometry.* A few of the quantities which are used and some of the simpler methods of measuring them will be explained.

The *luminous intensities* of different sources of light are measured in terms of *candle power.* The *international candle* is an arbitrary unit for measuring luminous intensity. A special group of incandescent lamps the candle power of which is known is maintained by the Bureau of Standards in conjunction with similar institutions abroad. Anyone may buy an incandescent lamp the candle power of which has been determined by a direct or an indirect comparison with these standards. This lamp can then be used as a standard to measure the candle power of other sources.

The unit for measuring *intensity of illumination* of any surface is so defined that

$$\text{Intensity of illumination} = \frac{\text{candle power of source}}{\text{square of distance from source}}.$$

When the distance is measured in feet, the intensity of illumination is expressed in *foot-candles.* For example, at a distance of 4 feet from a lamp having a luminous intensity of 24 candle power, the intensity of illumination is

$$\frac{24}{(4)^2} = 1.5 \text{ foot-candles.}$$

If distances are measured in meters, the intensity of illumination is expressed in *meter-candles*; that is, the candle power divided by the square of the distance in meters gives *meter-candles*.

For reading in daylight, the best intensity of illumination is about 100 foot-candles. The best amount to use depends on how well the light is distributed in the room, and on the absence of glare. With artificial light the best amount is usually from 4 to 10 foot-candles. If the distance from the light source is 3 feet, the candle power of the source must be 36 candles in order to give an illumination of 4 foot-candles; for

$$\frac{36}{(3)^2}=4.$$

Such a computation as this assumes that the book or paper is held *perpendicularly* to the direction the light is traveling. Simple reasoning shows that the intensity of illumination falls off rapidly as the book is turned away from this position.

The use of reflectors usually increases the illumination. With some types the increase is considerable. However, the ornamental shades used on many reading-lamps often decrease the illumination considerably.

562. The photometer. The measurement of the luminous intensity, or candle power, of a source of light is a matter of considerable importance. The principle that is involved can be stated in a simple way, as follows:

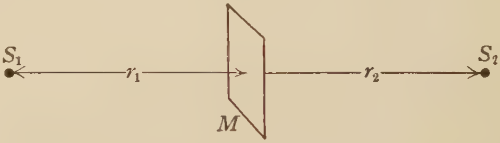


FIG. 355

Two sources of light are placed at two points from 1 to 2 meters apart (S_1 and S_2, Fig. 355). Let a screen M be placed between these, and moved to the right or left until a position is found where its two sides are equally illuminated by the sources S_1 and S_2. The eye is used for judging the equality of the illumination on the two sides of the screen. In different types of *photometer*, as this apparatus is called, different means are used for seeing the two sides simultaneously.

When the proper position of the screen is found, a simple equation holds, as can be readily demonstrated. Let I_1 be the candle power of S_1, and I_2 that of S_2.

The intensity of illumination on the left side of the screen is

$$\frac{I_1}{r_1^2},$$

where r_1 is the distance of the source S_1 from the screen. The intensity on the right side is

$$\frac{I_2}{r_2^2},$$

where r_2 is the distance of the source S_2 from the screen. When the illuminations of the two sides of the screen are equal,

$$\frac{I_1}{r_1^2} = \frac{I_2}{r_2^2}$$

or

$$\frac{I_1}{I_2} = \frac{r_1^2}{r_2^2}. \tag{5}$$

If the candle power of either S_1 or S_2 is known, the other can readily be computed from this equation.

We cannot measure directly the sensation produced on the retina of the eye. Yet we use this sensation in measuring the candle power of a source, for we use it to judge when two intensities of illumination are equal. Then, by the use of an arbitrary unit (that is, by assuming a value for some source), the candle power of another source is determined (the student should see why it is convenient to use an *arbitrary* unit,* an unusual thing in physics).

563. Theories of light. Apparently Plato and Aristotle thought that light was a property of the eye,—that something went out from the eye which enabled it to perceive objects; but common experience with a photographic camera, and many other kinds of observations, show us that the eye has nothing to do with the radiation of light, but that *energy* is in some way radiated from the *source* of light.

There are two well-known methods by which energy can be radiated. Suppose that a boat is floating in water a short distance off-shore. Energy can be transmitted to it in two different ways:

1. Energy can be transmitted to the boat by *wave-motion.* One standing on the shore might, by the aid of a board, produce

---

*See section 16 for a definition of an arbitrary unit.

waves in the water which would rock the boat and give it kinetic energy. (All types of wave-motion transmit energy (sect. 313).)

2. Energy can be transmitted to the boat by means of *projectiles*. A machine gun is an instrument for transmitting energy by this method. The cathode rays and the alpha and beta radiations from radioactive substances belong to this class.

The real tests of a theory are its ability to explain the known facts, and its ability to predict and thus to help us to discover new facts. These tests will be applied to the different theories of light.

The first theory of light which was developed in a truly scientific manner was that a luminous body shot off small particles which traveled at high speeds in straight lines. These small particles were called *corpuscles*, and the theory was known as the *corpuscular theory*. By the aid of this theory the laws then known of reflection, refraction, and many other phenomena were fully explained. Sir Isaac Newton was a strong advocate of this theory; and the weight of his authority, together with the ingenious explanations he devised, for a long time prevented the acceptance of the modern *wave theory*. Although Newton explained away many of the objections to the corpuscular theory, there remained some facts which were not explained satisfactorily. One of these facts is that light has a constant speed. It is very difficult to see why the corpuscules shot out from luminous bodies should all have the same speed, especially when it is recalled that a great many different types of bodies give out light. For example, there are stars which are a great many times larger than our sun—some so large that the gravitational attraction on the emitted corpuscles should be so great as to slow them up or even to stop them. The corpuscular theory also failed to explain satisfactorily certain well-known facts discussed under the headings *refraction*, *interference*, *diffraction*, and *polarization*. The student should watch for these and satisfy himself that the objections are really serious.

The modern *wave theory* of light states that light is a wave-motion of the transverse type (sect. 311). The radiation is not confined to one definite frequency or even to a few. Many frequencies are possible—practically an infinite number. In free space they all have the same speed; but in water or glass, and other transparent substances, different frequencies have different speeds. The details of this theory will be explained as needed in accounting for various observed facts.

Since light is propagated through the highest vacua, such as the space between heavenly bodies, it seemed necessary to assume some *medium* which could transmit this wave-motion. Obviously this medium could not be a gas nor any ordinary material substance, for the space between the heavenly bodies is devoid of such material. The hypothetical medium which filled all space and which transmitted light-waves was called the *ether*. The early theories supposed that light-waves were a mechanical vibration of this ether; but it is now known that they are *electromagnetic* waves, of the same type as those used in wireless telegraphy and telephony. The wave-lengths are very much shorter and the frequencies much higher than those used in wireless. Electromagnetic waves are moving electric and magnetic fields (sect. 555), but as yet it is not known what electric and magnetic fields really are.

According to the electromagnetic theory, waves of light are produced by oscillating electric charges. There is much indirect evidence which shows that the actual sources of light-waves are the oscillating electrons in the atoms of the luminous source. This theory of the production of light has received a great deal of attention and has stood the most severe tests. Not only does it explain known facts but it has predicted many facts which have subsequently been verified. Among the things which have been predicted by the electromagnetic theory is that light falling on a body should exert a *pressure* on it. The amount of this predicted pressure was very small; for example, the pressure produced in full sunlight on an absorbing surface should be about half a dyne (which is less than the weight of a milligram) per square meter of surface. This was so small an effect that for a long time physicists believed that it never would be measured or even detected; but in 1899 Lebedew in Russia detected it, and in 1901 Nichols and Hull in this country actually measured it and showed that its magnitude agreed with that predicted by the electromagnetic theory.

However, there are some important and serious difficulties connected with the wave theory of light. So far, an entirely satisfactory explanation of how, on the one hand, the energy of an oscillating electron is converted into wave energy or, on the other hand, of how an electron absorbs energy from a wave has not been found. The *quantum theory* has been developed, to explain some of the difficulties connected with the radiation and absorption of energy. A great deal of attention is now given to this new theory in advanced courses in physics.

564. The velocity of light. The earliest attempt to determine whether light has a finite velocity of propagation seems to have been made by Galileo. He stationed two sets of observers on two different hilltops. One group sent out a flash of light by uncovering a lantern. When the second group perceived this flash, they uncovered their lantern. The first set measured the interval between the time when they uncovered their lantern and the time when they received the return flash. Galileo knew that a serious error would be made on account of the time required for the observers to act—what we now call *reaction time.* To correct for this, he put the two sets of observers farther apart and repeated the experiment. But the speed of light was entirely too great to be detected in this manner. This is very clear when one realizes that light is known to travel about 186 miles in *one one-thousandth of a second*!

The earlier methods of measuring the velocity of light were based on astronomical methods. Later, terrestrial methods, far more accurate, were devised. A few typical methods, taken in historical order, will be explained in the following sections.

565. Römer's method. The first practical suggestion for determining the velocity of light was made about 1676 by Römer, a Danish astronomer, in connection with attempts to measure accurately the period (by *period* is meant the time required for a satellite to go entirely around the planet) of the satellites of the planet Jupiter. As a satellite moves around Jupiter it passes behind this planet, and hence is eclipsed. The most convenient way of measuring the period is to determine the interval between two of the times when the satellite disappears. After the period was determined in this way, it was assumed that the value obtained could be checked by predicting the time of future eclipses. This, the observers thought, should be a very simple method. But when this was done, it was found that eclipses did not occur on schedule time: sometimes they were ahead of time and sometimes belated. One suggestion for this discrepancy was that the motion of the satellite was irregular. But there seemed no way to account for such an irregular motion. Römer, working in Paris on observations made on what is called the first satellite, was the first to suggest that these irregularities were due to a finite velocity of the propagation of light. His explanation can be readily understood by the aid of the following:

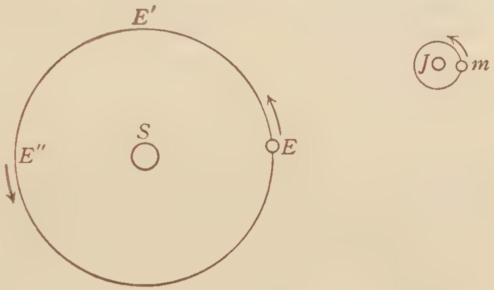


FIG. 356

In Fig. 356, S denotes the position of the sun, and $EE'E''$ the orbit of the earth; J is Jupiter, and m is one of the satellites which revolve around it. Jupiter revolves around the sun so slowly that its motion will be neglected. Let us suppose that the period of the satellite m has been carefully determined while the earth is near the position E. Knowing the value of the period (which in the case of the first satellite is about 42.5 hours) and the time of one eclipse (that is, the time when m disappeared behind Jupiter), the time of future eclipses can be predicted. Consider the predictions for the eclipses at a time about six months later, when the earth is near E''. As the earth is now much farther away from Jupiter, the eclipse will not be seen on schedule time, but will appear behind time by an amount equal to *the time required for the light to travel across the earth's orbit* from E to E''; but when the earth gets back to E, the eclipses should occur on schedule time. In this way the observed discrepancies in the predicted schedule of the eclipses were explained. By assuming this explanation to be the correct one, it was found that the time which is required for light to cross the earth's orbit—a distance of about 185,000,000 miles—is 16.5 minutes. This gives, for the velocity of light, about 186,000 miles per second.

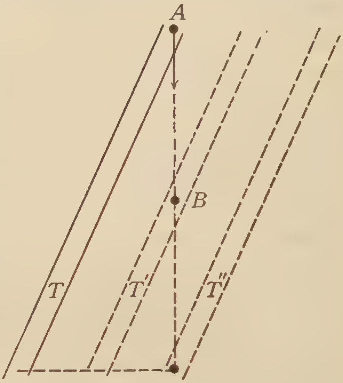


FIG. 357

566. Bradley and the aberration of light. An interesting method for measuring the velocity of light was developed by an Englishman, James Bradley, about 1728. The principle can be understood from the following analogous case: Suppose that a person is holding a long tube, T, in a direction shown in Fig. 357, and that a raindrop falling vertically has just entered the upper end at A. Let us suppose, further, that the person and the tube are carried along horizontally with a velocity so great that the raindrop does not strike the wall of the tube. If the person were looking up through the tube, the drop would not appear to be falling in a vertical direction but in a direction parallel to the inclined position of the tube.

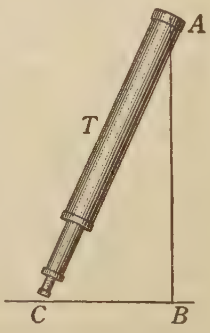


FIG. 358

The earth is moving around the sun with a mean velocity of about 18.5 miles per second. This velocity is so great that light from most of the stars appears to come from a slightly different direction. This effect is known as the *aberration* of light. If light from a star is traveling straight down, a telescope, T (Fig. 358), must be slightly inclined (exaggerated in the figure) so that the light can travel through it to the eye of the observer. It must be inclined at such an angle that CB is the distance the earth moves while light travels the distance AB. Since these distances are traveled over in equal times, we may say that

$$\frac{AB}{CB} = \frac{\text{velocity of light}}{\text{velocity of the earth}}.$$

The ratio AB/CB can be computed from the measured inclination of the telescope, and since the velocity of the earth is known, the velocity of light can be found by the aid of this equation.

If the earth always moved in the same direction, it would not be possible to measure the angle the telescope must be tipped, for we could not find out the true direction of a star. But the earth moves in an orbit with a period of twelve months; hence at times six months apart the direction of motion of the earth will be opposite. At one time the stars are apparently displaced in one direction, and six months later they are displaced an equal amount in the opposite direction; hence, from observations taken six months apart, the true direction of a star can be determined. There is no displacement in the case of the stars toward which we are moving nor in the case of those from which we are moving. The maximum displacement occurs for those stars seen at right angles to the motion of the earth. Even for this case the angular displacement is very small, recent results giving the angle as 20.49 seconds. Obviously this method cannot be used to obtain very accurate results for the velocity of light; nevertheless it is interesting from a historical point of view, and also because the aberration of light plays a very important part in certain rather difficult modern problems.

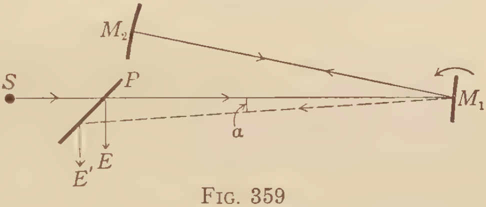


FIG. 359

567. Foucault's method. A terrestrial method for getting the speed of light was successfully used by the French physicist Foucault in 1850. The principles involved are relatively simple. Light from a source S (Fig. 359) passes through a plate of glass and falls on a mirror M_1, which can be rotated at a high rate of speed about an axis perpendicular to the plane of the figure. M_2 is a mirror so curved that light reflected to it from the mirror M_1 will be thrown directly back to M_1 over the same path and, if M_1 is not moved, from there back to the glass plate P, which will reflect part of the returning light to the eye at E. To understand the effect of the rotation of the mirror M_1, consider the instant that M_1 is in such a position that the light reflected from it will reach some point on M_2. From M_2 the light will be reflected directly back to M_1. But during the time the light has been traveling from M_1 to M_2 and back, the mirror M_1 will have moved through a small angle in the direction of the arrow. Hence the beam reflected back from it will pass back over a different path (as indicated by the broken line in the figure) to some point E'.

In the figure much has been omitted; for example, the reflected beams at E and E' were brought to a focus in the field of a micrometer microscope, so that the distance between the image at E formed when M_1 was stationary, and that at E' when M_1 was in rotation, could be accurately measured. From this measurement and from the known distance to the mirror M_1 the angle α of Fig. 359 could be computed. It can be shown by simple geometry and by the law of reflection (sect. 569) that when a mirror is turned through an angle, the beam reflected from it will be turned through twice that angle.* Hence, if the light has been deflected through an angle α by the motion of the mirror, the mirror must have turned through an angle $\frac{1}{2}\alpha$ while the light traveled from M_1 to M_2 and back. It is easy to compute the time required for M_1 to rotate through the angle $\frac{1}{2}\alpha$ when both the angle and the rotational speed of M_1 are known. The final result for the speed of light was obtained by dividing twice the distance M_1M_2 by the time required for the mirror to turn through the small angle $\frac{1}{2}\alpha$.

The mirror M_1 was revolved at the very high speed of eight hundred revolutions per second. The distance M_1M_2 at first was about 4 meters. Later a distance equivalent to about 20 meters was used. In both cases the shift of the light from E to E' was very small—less than 1 millimeter.

While Foucault's results were not accurate, he did one thing which at that time represented a very important experiment. He placed a tube of water between M_1 and M_2, so that the light traveled through water. In this way he was able to show that *the*

---

* The student can easily test this. Hold a mirror so that the light is reflected straight back to the source. Then turn the mirror through an angle of 45 degrees. The reflected light beam should be changed by 90 degrees.

velocity of light in water is less than that in air. This result gave the deathblow to the corpuscular theory of light; for, as will be explained in section 589, the corpuscular theory demands that light travel faster in water than in air, while the wave theory demands that it travel slower.

Foucault's method was greatly improved by Michelson in 1879–1882 and by Newcomb in 1882. Probably the most accurate measurement of the velocity of light is being made by Michelson in California. Light is reflected by a rotating mirror on Mt. Wilson to a stationary mirror on Mt. San Antonio, 22 miles away, and reflected back. As a result of preliminary observations during the summer of 1924 he obtained for the velocity of light in vacuo 299,820 kilometers per second, or 186,300 miles per second.

PROBLEMS

1. A 25-watt lamp, vacuum type, is rated at 1.1 watts per candle power. What intensity of illumination in foot-candles is produced at a distance of 3 ft. from the lamp?

2. A 16-candle-power electric lamp is 3 ft. from a book. How far from the book must a 300-candle-power lamp be placed if it is to produce the same intensity of illumination?

3. What candle power is required to produce an illumination of 6 foot-candles at a distance of 4 ft. from the source?

4. An incandescent lamp, gas-filled type, is rated at 0.70 watt per candle. What size of lamp, rated in watts, is needed to produce, at a distance of 3 ft., an illumination of 8 foot-candles?

5. In a photometer experiment it was found that two lamps produced the same intensity of illumination at a point 120 cm. from the first lamp and 80 cm. from the second. If the candle power of the first lamp was 24, what was the candle power of the second?

CHAPTER XLI

REFLECTION

Diffuse and regular reflection, 568. Law of regular reflection, 569. Huygens' principle, 570. Proof of the law of regular reflection, 571. The ray and wave methods, 572. The image in a plane mirror, 573. Virtual and real images, 574. The concave mirror, 575. The principal focus, 576. Conjugate foci, 577. Wave construction for a concave mirror, 578. The graphical construction of real images, 579. Virtual images formed by a concave mirror, 580. Law of size of image, 581. Formula for position of object and image, 582. Derivation of the formula, 583. Principal focus of a convex mirror, 584. Graphical construction of images formed by convex mirror, 585. Formula for a convex mirror, 586. Spherical aberration, 587.

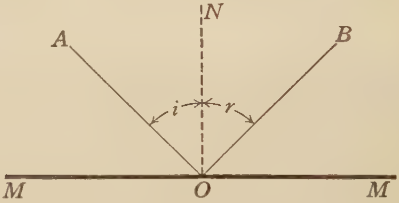


FIG. 360

568. Diffuse and regular reflection. It is convenient to distinguish between two types of reflection: *diffuse* and *regular*. When a sheet of white paper is placed in a beam of light, the light is reflected or scattered in all directions. It is said to be diffusely reflected. The light from the page of this book and from all or nearly all the objects in an ordinary room reaches the eye by diffuse reflection; but when a beam of light falls on a mirror, it is all, or practically all, reflected in some one direction. *Regular* reflection occurs only from smooth surfaces, such as those of glass, mercury, water, and polished metals.

569. Law of regular reflection. In Fig. 360, *MM* is the trace of a mirror placed perpendicularly to the plane of the figure, *AO* the direction the incident light is traveling, *OB* the direction of the regularly reflected light, and *ON* a line perpendicular, or normal, to the mirror *MM*. The *angle of incidence* is the angle *AON*, and the *angle of reflection* is the angle *NOB*.

In the case of regular reflection the angle of reflection is equal to the angle of incidence, and the two angles lie in the same plane. It is rather a simple matter to prove by direct experiment that this law is true. However, a proof based on the wave theory of light will be given. But it will be necessary first to explain a principle that we must use.

570. Huygens' principle. A method due to Huygens affords a simple means of finding the change in the position of any wave. Let the given wave be AB (Fig. 361). It is required to find the position of this wave at some subsequent time. A number of points on this wave—P_1, P_2, etc.—are chosen. From each of these is drawn an arc of a circle the radius of which has the same length in all cases. Huygens' principle states that the new position of the wave is represented by a line which is tangent to these arcs—the curve $A'B'$ in the figure.

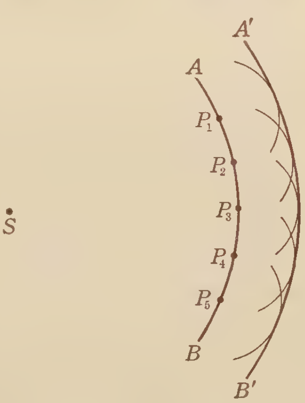


FIG. 361

The principle may be regarded merely as a geometrical one. But there is a deeper meaning attached to it. This may be summarized as follows: At each point of an advancing wave the medium is disturbed, and a disturbance is propagated from this point in all directions. P_1, P_2, P_3, etc., in Fig. 361, are centers of disturbance. But these disturbances interfere (sect. 329 and Chapter XLVI) with one another in such a way that the resultant disturbance is found only along the surface which is tangent to the wavelets sent out from each of these centers.

571. Proof of the law of regular reflection. In Fig. 362, AB represents an incident wave, and AC the mirror. If the mirror did not interfere, the wave AB would, after a certain time, reach a position DC. A number of points M_1, M_2, and M_3 are taken on the surface of the mirror. At different times these points become centers of disturbances which, if the mirror were not there, would reach DC at the same time and, in accordance with

Huygens' principle, form the wave DC. An arc of a circle with a radius equal to AD is drawn with A as the center, another with M_1 as the center and with a radius equal to the distance of M_1 from CD (that is, the distance M_1N_1), another with M_2 as a center and with a radius equal to M_2N_2, and another with M_3 as a center and with a radius of M_3N_3. Since the line FC is tangent to these arcs, it gives, by Huygens' principle, the position of the reflected wave.

To show that the angle of reflection is equal to the angle of incidence, it is necessary merely to show that the two lines BC and AF, which represent the directions the incident and reflected waves are traveling, make equal angles with the mirror AC. Since the line AF is equal to AD by construction, and since AD is equal to BC, AF is equal to BC. The side AC is common to the triangles ABC and AFC. Since the two triangles have two sides equal, and since they both contain right angles, they must be equal; hence the angle BCA is equal to the angle FAC. But the angle of incidence is not the angle the line BC makes with the reflecting surface: it is the angle that BC makes with the *normal* to the surface. However, it should be entirely clear to the student that if the angles BCA and FAC are equal, the angles of incidence and reflection must be equal.

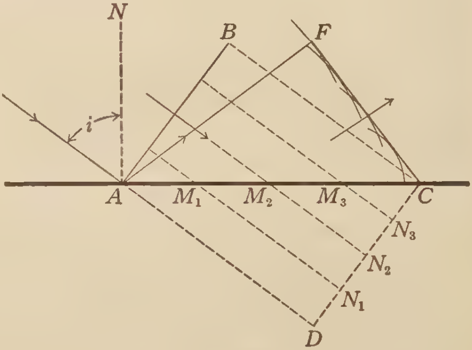


FIG. 362

It should be obvious that since this proof applies to any type of wave-motion, the law of regular reflection applies not only to light-waves but to all kinds of waves.

572. The ray and wave methods. Since light usually travels in straight lines, it is often simpler to deal only with a line which indicates the direction the light is traveling. Such a line is often called a *ray* of light, but the term has no other physical significance. Light does not consist of bundles of rays: it is a wave-motion. While the ray method is often very simple, it has the disadvantage of tending to conceal the true mode of propagation. In the next section a simple problem will be worked out by both the ray and the wave method.

573. The image in a plane mirror. In the use of the ray method, lines, or rays, are drawn in the direction the light is traveling. In Fig. 363, O is the source of light, the line OP_1 is one ray, and OP_2 is another. The line P_1R_1 is drawn so that it makes an angle with the normal to the surface at P_1 equal to the angle made by OP_1 with the same normal. P_1R_1 then represents the ray reflected at P_1. In a similar manner the line P_2R_2 represents the ray reflected at P_2. The two rays P_1R_1 and P_2R_2 appear to come from the point I, which is located by continuing these lines back through the mirror. I is called the *image* of O. If a large number of rays were drawn from O, the reflected rays would all appear to come from the image I.

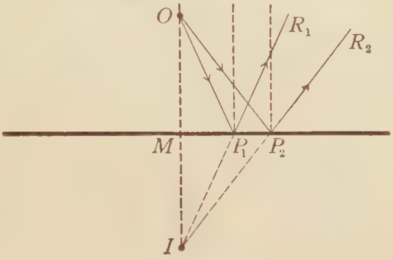


FIG. 363

It will now be proved (1) that the line OI joining the object and the image is perpendicular to the mirror, and (2) that the image is just as far behind the mirror as the object is in front.

1. First draw a line from O perpendicular to the mirror (Fig. 364), and then produce the ray PR backward until it intersects the perpendicular line at O'. The line PN is also drawn perpendicular to the mirror. It follows from the construction and from the law of reflection that the angles OPN, NPR, MOP, and $MO'P$ are all equal. It then follows, rather simply, that the triangle MOP is equal to $MO'P$; hence OM is equal to MO'.

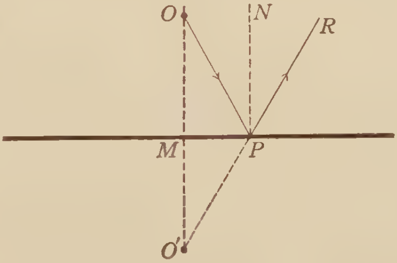


FIG. 364

But this is true *no matter where the point P is taken on the surface.* If another ray is drawn from O to some point on the surface, and if the reflected ray is drawn backward until it cuts the line OM, it will pass through the point O'; hence O' must be the image of the object. Since O' is a point on the perpendicular line OM, the line connecting the image and the object is perpendicular to the surface of the mirror.

2. The second fact, that the image lies just as far behind the mirror as the object is in front, is proved as soon as it is shown that the point O' is the position of the image; for it has been shown that OM is equal to MO'.

It is not necessary to go into great detail to explain the wave method, for it can be understood by the aid of Fig. 365. The source of the waves is at O. If the reflecting surface MN did not interfere, a wave would reach the position ABC. But the wave is turned back by the surface and travels just as far back as it would otherwise go forward. The method of finding the reflected wave B' can be seen in the figure. Several points on the mirror are chosen, and circles are drawn which are tangent to the wave ABC. At some time each of the centers of these circles is a center of disturbance, and the wavelet emitted travels upward the same distance it would go downward if the mirror MN were not there. The reflected wave B' is drawn tangent to these circles. The wave B' in traveling back from the surface appears to come from the geometrical center of the wave, the point I, which is therefore the image of O. From symmetry it can be seen that the center I of the wave B' must be just as far below the plane as the source O (the center of the wave B) is above it.

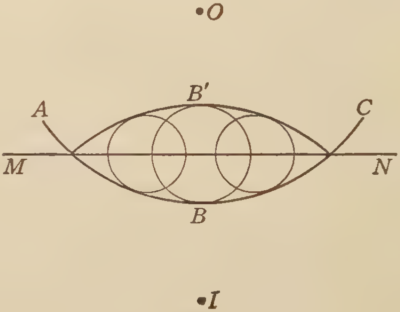


FIG. 365

574. Virtual and real images. The images formed in plane mirrors are called *virtual* images, to distinguish them from *real* images, where the light actually reaches the image. The images formed on screens—for example, those of lantern slides, motion pictures, etc.—are real images. When one sees his own image in a plane mirror, the light does not go to this image—it only seems to come from this place. Hence it is a virtual image.

575. The concave mirror. A concave spherical mirror will, under certain conditions, produce *real* images of objects; for example, if a candle flame or an incandescent lamp is placed just a little below the *center of curvature C* of the mirror (Fig. 366), a real, inverted image will be formed just above C. An eye in front of the mirror and some distance away (to the right in the figure) will see this image apparently hanging in mid-air. If a screen of paper or ground glass is placed there, the image will be pictured on the screen.

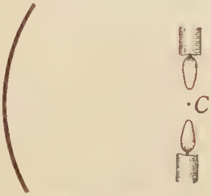


FIG. 366

576. The principal focus. As will appear later, real images can be formed at many different distances from a concave mirror if the source is placed in different positions. There is one position of the image which is called the *principal focus.*

The principal focus is that point where the image is found when the source is at a great distance away.

When the source is small and at a great distance, the rays from it are practically parallel; hence the principal focus may be defined as the *point where parallel rays are focused.*

It will now be proved that the principal focus is halfway between the center of curvature and the mirror. In Fig. 367, SP, which is drawn parallel to the axis MC, where C is the center of curvature, represents one of the incident rays, and PF represents a reflected ray. Since the line PC is a radius of the sphere, it must be normal to the mirror at the point P. According to the law of reflection the angle SPC must be equal to the angle CPF. But the angle SPC is equal also to the angle PCF, because the lines PS and MC are parallel;

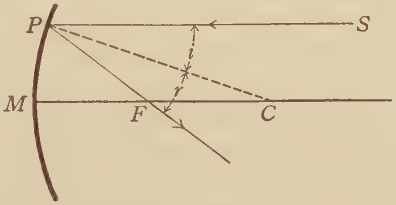


FIG. 367

hence the angles *CPF* and *PCF* are equal, and the triangle *PFC* is an isosceles triangle. Hence *PF* is equal to *FC*. In the use of spherical mirrors sharp images can be formed only when the rays are reflected from points close to the center *M*; hence in actual practice the point *P* is near *M*. When it is, the line *PF* is so nearly equal to *MF* that we may take them as equal. Since *PF* is equal to *FC*, *MF* is approximately equal to *FC*, or the point *F* is very nearly halfway between the mirror and the center of curvature.

The distance *MF* is called the *focal length* of the mirror and will be denoted by *f*.

When the source of light is placed at the principal focus, the reflected rays will be parallel; that is, if *FP* (Fig. 367) is a ray leaving *F*, it will be reflected parallel to the axis. This is one instance of a general principle of the interchangeability of the source and the image (explained in section 577).

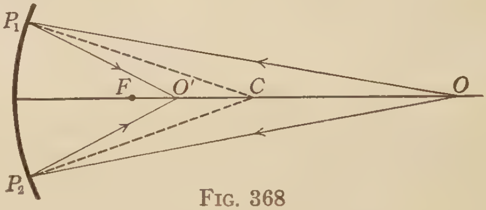


FIG. 368

577. Conjugate foci. If a source of light is placed at *O* (Fig. 368), its image will be formed at *O'*. The position of the image can be found as follows: A ray from *O* is reflected at P_1 so that the angle of reflection CP_1O' will be equal to the angle OP_1C. In a similar way the ray OP_2 is reflected to *O'*. The intersection of the two rays P_1O' and P_2O' locates the image at *O'*. When the object is placed at *O'*, its image will be formed at *O*. For in the figure, $O'P_1$ may be regarded as the incident ray and P_1O the reflected ray; similarly with $O'P_2$ and P_2O. The two points *O* and *O'* are called *conjugate foci*. Two points are said to be conjugate whenever the object and the image can be interchanged; that is, no matter which point is taken as the object, the other point will be the image.

578. Wave construction for a concave mirror. In Fig. 369, *C* is the center of curvature of the mirror *M*, and *O* and *O'* are two conjugate points—one the center of the arc P_1, the other the center of the arc P_2. If the source is at O, P_1 is a position a wave would have reached if the mirror had not been present; but the incident wave is reflected by the mirror and, instead of going to P_1, goes to P_2. The position of the reflected wave is found by Huygens' principle, a few wavelets being shown in the figure. The reflected wave will *converge* toward its center of curvature; that is, toward O'. The center of a converging wave is always a real image.

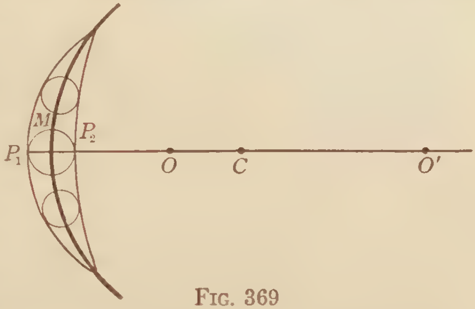


FIG. 369

579. The graphical construction of real images. The position of the images formed by a concave mirror can be located by a simple graphical method. Let C of Fig. 370 be the center of curvature of the mirror, and F the principal focus, which is halfway between C and the mirror. Let O be a point on the object the image of which we wish to locate. A ray OB is drawn

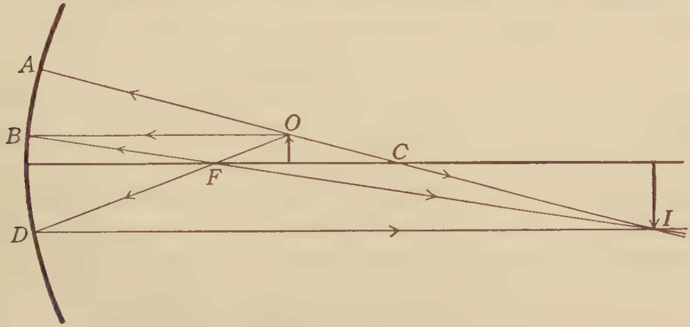


FIG. 370. Real image formed by concave mirror

parallel to the axis. From the definition of the principal focus this parallel ray must, after reflection, pass through the point F, as shown. This gives the path of one ray after reflection. Draw another ray, OA, in the line COA, which passes through the center of curvature C. As this falls normally on the mirror, it will be reflected directly back through C. This gives the path of a second ray. A ray OFD passing through the principal focus F will be reflected parallel to the axis, as shown. In this way the path of a third ray is obtained. The intersection of the three reflected rays locates I, the position of the image of the point O. Obviously it is not necessary to draw three rays; one may take his choice as to which two shall be used.

Usually it is sufficient to find the image of one point of an object. For example, in Fig. 370 the position of the image of the head of the arrow, O, is found to be at I. That is enough to locate the image. This graphical method gives the location of an image quickly, and with sufficient accuracy for many purposes.

By the aid of such drawings as Fig. 370 the student should see that the real image of an object is always *inverted.*

580. Virtual images formed by a concave mirror. The two positions O and I in Fig. 370 are conjugate foci. If the object is at I, the image is at O; or, if the object is at O, the image is at I. Suppose the object is at I, and we move it toward the mirror. In what direction does the image move? It moves away from the mirror. When the moving object reaches the center of curvature C, the image will meet it there. Why? When the object reaches the position O, the image will be at I. When the object reaches F, the principal focus, the reflected rays will form parallel beams, and the image may be said to be at an infinite distance away. What happens when the object is moved inside F? This question will be answered by finding the position of the image by graphical construction.

In Fig. 371, F is the position of the principal focus, C the center of curvature, and O a point on the object. The ray OB, which is drawn parallel to the axis, is reflected back through F toward B'. The ray OD, traveling in the same direction as if it had come from the principal focus F, is reflected parallel to the axis, in the direction of D'. The ray OA, falling normally on the mirror, is reflected straight back through C, the center of curvature. The reflected rays DD', AC, and BB' are *diverging* and will never meet. However, they *appear* to come from the point I, which is therefore a *virtual* image of O. In this case the image is *erect*, *virtual*, and *magnified.*

Hence, when the object is inside the principal focus, the image is behind the mirror and is erect, virtual, and magnified.

581. Law of size of image. A very simple law gives a relation between the sizes of object and image and their distances from the mirror. If p is the distance of the object from the mirror, and q the distance of the image from the mirror, the law is as follows:

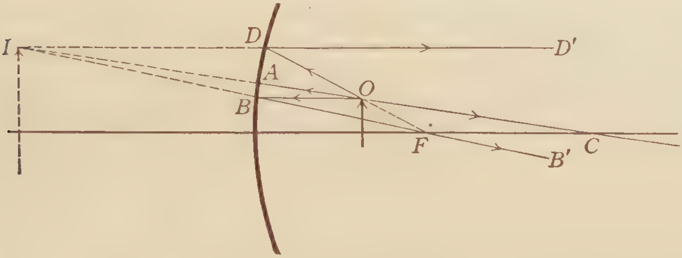


Fig. 371. Virtual image formed by concave mirror

$$\frac{\text{Size of object}}{\text{Size of image}} = \frac{p}{q}. \qquad (6)$$

It is relatively easy to derive this law. Let OA (Fig. 372) be half the size of the object and let BI be half the size of the image. The ray OM after reflection will pass through I, the image of O. Since the axis BCM is perpendicular, or normal, to the mirror at M, the angle OMA is equal to the angle BMI.

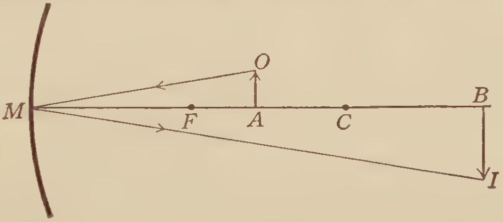


Fig. 372

Since the two triangles OMA and BMI contain right angles and have their acute angles equal, they must be similar triangles. Therefore the ratios of corresponding sides must be equal, or

$$\frac{OA}{BI} = \frac{MA}{MB} = \frac{p}{q}.$$

But *OA* is half the size of the object, and *BI* half the size of the image. Obviously the ratio of their half-sizes must be equal to the ratio of their sizes; equation (6) follows at once.

It is not difficult to show that the same relation holds when the image is virtual.

582. Formula for position of object and image. There is always a definite numerical relation between the radius of curvature of the mirror R, the distance of the source from the mirror p, and the distance of the image from the mirror q. The equation stating this relation is

$$\frac{1}{p}+\frac{1}{q}=\frac{2}{R}. \tag{7}$$

By the aid of this formula, when any two of these quantities are known, the third can be computed.

For example, suppose that the radius R of the mirror is 30 centimeters, and that the object is 20 centimeters from the mirror (that is, $p=20$); then

$$\frac{1}{20}+\frac{1}{q}=\frac{2}{30},$$

$$\frac{1}{q}=\frac{2}{30}-\frac{1}{20}=\frac{1}{60}.$$

Therefore $$q=60 \text{ cm}.$$

Three special cases will be considered:

1. When $p=R/2$ (that is, when the object is at the principal focus), we have

$$\frac{2}{R}+\frac{1}{q}=\frac{2}{R},$$

or $$\frac{1}{q}=0.$$

Hence q is an infinitely large number; the image is at an infinite distance from the mirror.

2. When p is very great (that is, when the object is at a great distance away), $1/p$ is practically zero; then

$$0+\frac{1}{q}=\frac{2}{R},$$

or $$q=\frac{R}{2}.$$

Hence the image is formed at the principal focus.

3. Equation (7) may be written

$$\frac{1}{q}=\frac{2}{R}-\frac{1}{p}. \tag{8}$$

When the object is inside the principal focus (that is, when p is less than $R/2$), $1/p$ will be greater than $2/R$. Hence, from equation (8), $1/q$, and therefore q, will be a *negative* quantity. In this case the image is behind the mirror and is a *virtual* image. We may conclude that whenever q is a negative quantity, the negative sign means that the distance is to be measured from the back of the mirror.

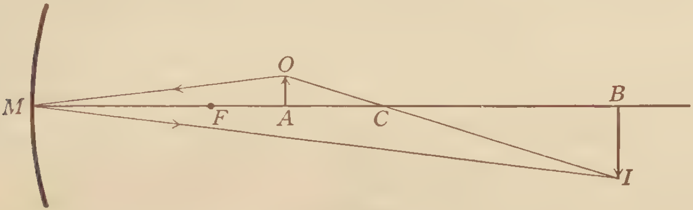


Fig. 373

A numerical problem will illustrate this last case. Suppose that the radius of the mirror is 30 centimeters, and that p, the distance of the object from the mirror, is 10 centimeters. Using equation (8), we obtain

$$\frac{1}{q}=\frac{2}{30}-\frac{1}{10}=\frac{2}{30}-\frac{3}{30},$$

or

$$\frac{1}{q}=-\frac{1}{30}.$$

Hence

$$q=-30\text{ cm.}$$

The image is a virtual one, 30 centimeters behind the mirror.

583. Derivation of the formula. The formula given in equation (7) will now be derived. In Fig. 373, OA is the object and BI is the image. A line drawn from O to I must pass through C, the center of curvature. To see that this is true, imagine a ray drawn from O toward the mirror,—a prolongation of the line CO. It will strike the mirror normally, and be reflected straight back through

O and C to the image which is at the point I. The angle OMA must be equal to the angle BMI; for a ray OM after reflection must pass through I, and the angle of incidence must be equal to the angle of reflection. The triangles OMA and BMI are then similar, as are the triangles OAC and CIB. Taking the corresponding sides of the similar triangles,

$$\frac{OA}{BI}=\frac{AC}{CB}$$

and

$$\frac{OA}{BI}=\frac{MA}{MB}.$$

Hence

$$\frac{AC}{CB}=\frac{MA}{MB}. \tag{9}$$

But from the figure it is easily seen, in terms of the notation we are using, that

$$MA=p,$$
$$MB=q,$$
$$AC=R-p,$$

and

$$CB=q-R,$$

Substituting these values in equation (9), we have

$$\frac{R-p}{q-R}=\frac{p}{q}.$$

Clearing of fractions,

$$Rq-pq=pq-pR,$$
$$Rq+pR=2pq.$$

Dividing this equation by pqR, we have equation (7),

$$\frac{1}{p}+\frac{1}{q}=\frac{2}{R},$$

the desired equation.

584. Principal focus of a convex mirror. When a parallel beam falls on a convex mirror, it is converted into a beam which is *diverging* from the point F (Fig. 374). F is the principal focus, and is a virtual focus.

It can be shown, by a method similar to that of section 576, that this principal focus is halfway between the center of curvature and the mirror.

585. Graphical construction of images formed by convex mirror. In Fig. 375, *C* and *F* are the center of curvature and the principal focus, respectively, of a convex mirror. From this figure it is relatively easy to understand the method of locating the image of *O*. A ray *OA* drawn parallel to the axis is reflected as if it came from *F*, the principal focus. A ray *OB* drawn in the direction of *OC* will be reflected back in the same line, because it strikes the mirror normally. A ray *OD* drawn in the direction of *OF* is reflected parallel to the axis. These three reflected rays, *AM*, *BN*, and *DP*, are diverging, and will never meet. However, they appear to come from the image of *O*, which is located by continuing these lines back through the mirror.

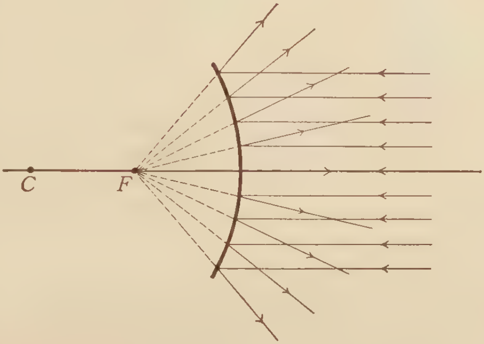


FIG. 374

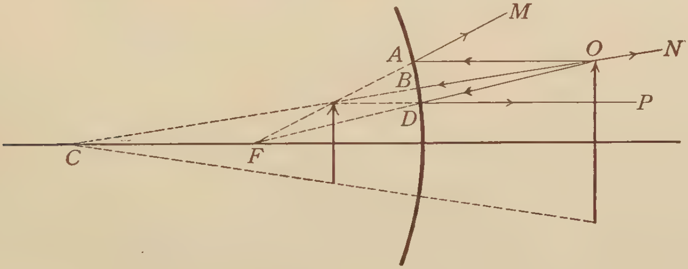


FIG. 375. Virtual image formed by convex mirror

Whatever may be the distance of the object from the mirror, the image formed by the convex mirror is always virtual. This can be proved by actual trial with a mirror or by constructing a number of figures similar to Fig. 375. When the object is very far away, the image is formed at *F*; as the object approaches the mirror, the image moves to meet it. For all positions of the object at finite distances the image lies somewhere between F and the mirror. While both real and virtual images can be formed by a concave mirror, *only virtual images are formed by a convex mirror.*

586. Formula for a convex mirror. In section 582 it was pointed out that a quantity with a *negative* sign means that the distance should be measured behind the mirror. Thus, when the image is virtual, the numerical value of q is negative. If this rule is observed, the formula (equation (7))

$$\frac{1}{p}+\frac{1}{q}=\frac{2}{R}$$

can be applied to a convex mirror. In using this formula the radius of curvature R must be treated as a *negative* quantity, for the center of curvature lies back of the mirror.

As an example take the case of a convex mirror, with a radius of 40 centimeters, when the object is 20 centimeters away.

Then $R=-40$

and $p=20.$

From equation (7)

$$\frac{1}{20}+\frac{1}{q}=-\frac{2}{40}=-\frac{1}{20},$$

$$\frac{1}{q}=-\frac{1}{20}-\frac{1}{20}=-\frac{1}{10},$$

or $q=-10$ cm.

Hence the image lies 10 centimeters back of the mirror.

It is easy to see from equation (7) that q, the distance of the image from the convex mirror, must always be negative,—that the image must lie behind the mirror. This equation may be written in the form (equation (8))

$$\frac{1}{q}=\frac{2}{R}-\frac{1}{p}.$$

In the case of the convex mirror, R is always negative; hence the right-hand side of this equation is always negative. Therefore q is always negative, and the image is always a virtual one.

587. Spherical aberration. In the proof given in section 576, it was assumed that reflection took place only from a small part of the mirror near the center. It is true that sharp images can be formed by spherical mirrors only when a small portion of a mirror is used. The blurring which occurs when a relatively large portion is used is called *spherical aberration.*

The effect of spherical aberration is readily demonstrated. An image of some bright, sharply defined object, such as the filament of an incandescent lamp, is formed on a screen by a concave mirror of moderate size. The mirror is moved until the sharpest image is obtained. If the mirror is now held still, and it is covered with a piece of cardboard with a hole in it, the image becomes appreciably sharper. The remedy for spherical aberration is to cover up all the mirror except a relatively small portion.

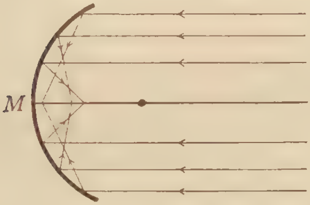


Fig. 376

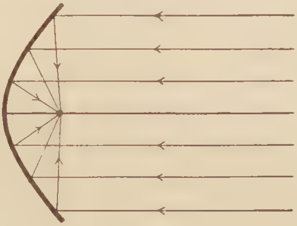


Fig. 377

In Fig. 376 is shown the reflection of a parallel beam of light when a spherical mirror of large "aperture" is used. Only those rays reflected from a region lying near the center M converge to the principal focus; but when the mirror is parabolic, all the parallel beam converges to the principal focus (Fig. 377). If a *small* source is placed at the principal focus of the parabolic mirror, the reflected beam is a parallel beam. But parabolic mirrors are useful only in the case of parallel beams. If it is desired to form an accurate image of a small object near at hand by reflection from a large mirror, the surface of the mirror must be a portion of an ellipsoid (Fig. 378) with the object and image, F_1 and F_2, at the foci. It is one of the properties of the ellipsoid that light starting from one focus F_2 will be converged to F_1, no matter what part of the surface reflects the light. But such a surface does not give sharp images for any other position of the object. Although the spherical mirror does have the defect of spherical aberration, it is better than any other form when it is desired for all sorts of use; that is, when the object is sometimes at one distance and sometimes at another. When the mirror is small compared with its radius of curvature, there is no practical difference in the shapes of the parabolic, the ellipsoidal, and the spherical mirror; hence spherical mirrors of small aperture are accurate enough for all practical purposes.

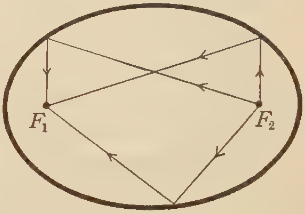


Fig. 378

PROBLEMS

1. Find the position of the image formed by a concave mirror of radius 30 cm. when the object is placed (*a*) 60 cm. from the mirror; (*b*) 20 cm. from the mirror; (*c*) 10 cm. from the mirror.

2. Find by graphical construction the position of the image for the case of a concave mirror, radius of curvature 10 in., with an object 8 in. from the mirror.

3. Compute the position of the image formed by a convex mirror of radius 20 cm. when the object is (*a*) 40 cm. away; (*b*) 10 cm. away.

4. Find the positions of the images of Problem 3 by graphical methods.

5. A concave mirror has a radius of curvature of 20 cm. An object 4 cm. high is placed 5 cm. from the mirror. Calculate (*a*) the position and (*b*) the size of the image.

6. An object 10 cm. high is placed 20 cm. from a convex mirror of radius 30 cm. Calculate (*a*) the position and (*b*) the size of the image.

7. It is desired to produce a real image 2 ft. high, of an object 4 in. high, by means of a concave mirror with a radius of 10 in. Where must the object be placed?

8. A concave mirror for shaving is designed to give a virtual image of one's face, at a distance of 15 in. from the eyes, with a magnification of 2. What is the radius of curvature?

CHAPTER XLII

REFRACTION

Refraction, 588. The cause of refraction, 589. The sine of an angle, 590. Snell's law of refraction, 591. Proof of the sine law by the wave theory, 592. Total reflection; the critical angle, 593. Examples of total reflection, 594. The relative index and the absolute index, 595. The prism, 596. Dispersion, 597. Newton's experiment with colors, 598. Atmospheric refraction of light, 599. The mirage, 600.

588. Refraction. When light passes from one medium into another, the rays are usually bent at the surface separating the two media. A simple method of illustrating this bending, or *refraction*, of light is to use a small tank, or aquarium, with sides and ends of glass. If a little fluorescein is added to the water in the tank, the path of the light through the water will be luminous. Fig. 379 indicates the appearance when a narrow beam of light is incident on the surface at an angle of about 45 degrees. As the beam *enters the water* it is *bent toward the normal*, or perpendicular, to the surface; that is, the angle of incidence is greater than the angle of refraction (the angle of refraction is the angle the refracted beam makes with a normal to the surface). As illustrated in the figure, the beam falls on a mirror M lying on the bottom of the tank, and is reflected back to the surface and again refracted. When the beam passes from *water to air*, it is *bent away from the normal*; that is, the angle of incidence is less than the angle of refraction.

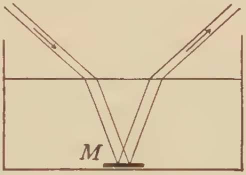


FIG. 379

When one looks at objects lying on the bottom of shallow water, they appear to be nearer the surface than they really are. In Fig. 380 let O_1 be a stone or some other object. A ray of light from the stone will reach the eye at E by the path shown. Because of the bending, the eye sees a virtual image of O_1 at I_1. An object O_2, a little farther away, appears to be at I_2; another, at O_3, appears to be at I_3. It is very easy to observe in shallow water of uniform depth that the farther parts look shallower. This effect is shown in the figure by the fact that I_2 is nearer the surface than I_1, and I_3 nearer than I_2. In clear water one may actually wade into much deeper water thinking that he is going into more shallow water.

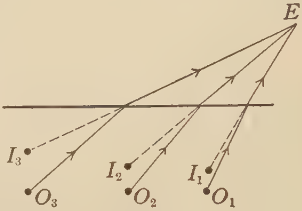


FIG. 380

589. The cause of refraction. In the history of optics there have been two important theories, each of which has been used to explain refraction: the *Newtonian corpuscular theory* and the *wave theory*. It so happens that these two theories lead to opposite conclusions in explaining refraction. Hence it becomes possible, as we shall see, to choose between them.

The corpuscular theory. According to the corpuscular theory, light consists of small particles, or corpuscles, which are shot off from a source at very high speeds. To explain refraction it seems necessary to suppose that these particles are attracted by matter, and, as they approach a denser medium, are given an acceleration toward the medium they are entering. In Fig. 381 is shown the path of a corpuscle as it enters a denser medium. The vectors A and B represent the tangential and perpendicular components of the velocity of the light-corpuscle before it enters the medium. A' and B' represent the components of the velocity of the light in the denser medium. The attraction of the medium for the corpuscle increases the perpendicular component of the velocity; hence B' is greater than B. But this attraction does not change the tangential component, so that A' is the same as A. This explanation leads to two conclusions:

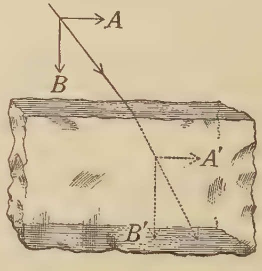


FIG. 381

1. Since the perpendicular velocity B' is greater than B, the resultant of A' and B' is a velocity which is different in direction

from the resultant of A and B; hence a *change in the direction of the ray* (refraction) must take place.

2. The resultant of A' and B' is greater in magnitude than the resultant of A and B; hence *the velocity in the denser medium is greater than it is outside.* This latter conclusion flatly contradicts that given by the wave theory, as we shall see later.

When a corpuscle emerges from the denser medium, the attraction causes a retardation of the corpuscle; hence it decreases the velocity and bends the path, as shown in Fig. 382. As before, the component tangential to the surface is not changed; that is, A' is equal to A. But owing to the attraction of the matter of the medium, B is less than B'.

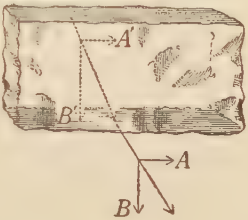


FIG. 382

The wave theory. According to the wave theory the refraction as light passes from air to water or to glass is caused by a *change in the velocity* of the waves. To account for the observed direction of the bending, it is necessary to assume that *light travels more slowly in water or glass than it does in air.* In Fig. 383 is shown the effect of a change in the velocity of waves as they enter a medium where the velocity is less (see also section 323). In general, waves travel in directions perpendicular to themselves. Hence in the case shown in the figure the arrows, drawn perpendicular to the waves, indicate the directions in which the light is traveling.

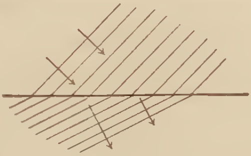


FIG. 383

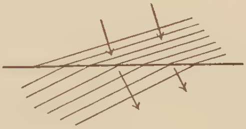


FIG. 384

When light travels from one medium into another where its velocity is greater, it will be bent away from the normal (Fig. 384).

We thus see that in order to explain refraction the *corpuscular* theory states that the velocity of light in water is *greater* than that in air, while the *wave* theory states that the velocity in water must be *less* than that in air. In 1850 Foucault announced that he had proved by direct experiment (sect. 567) that the velocity of light is less in water than in air. As pointed out in that section, this was a crucial experiment, and a deathblow to the corpuscular theory. It was not this alone, however, but this together with other observed phenomena in interference, diffraction, and polarization (Chapters XLVI and XLVII) that caused the abandonment of the corpuscular theory. Foucault's experiment was "the last straw."

More recent, accurate measurements of the velocity in water and in a few other transparent substances show that the actual change in the velocity is the proper numerical value to account for the amount of refraction observed. There is now no doubt that the wave theory gives a correct explanation of refraction.

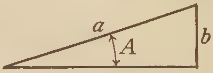


FIG. 385

590. The sine of an angle. Before learning the quantitative law of refraction it will be necessary for each student to be certain that he understands what is meant by the *sine* of an angle. Let the angle A of the right-angled triangle of Fig. 385 be the angle the sine of which will be defined. The *sine* of the angle A is defined as *the ratio of the side b to the hypotenuse a*, or

$$\text{Sine } A = \frac{b}{a}.$$

The magnitude of the sine of an angle depends only on the size of the angle; hence it may be used as a measure of the angle.

Angles may be measured in a number of different ways: by the well-known units *degrees*, *minutes*, and *seconds*, or by *radians*, or by the *ratio* of the *sides* of the right-angled triangle (for example, by the *sine* of the angle). Often it is more convenient to state the angle in terms of its sine or some other ratio of the sides of the triangle (as the *cosine* or the *tangent*) than in degrees. In the case of work involving algebraic equations angles are commonly stated in terms of their sines, or of one of the other ratios of sides of the triangle.

The following brief table is given to enable the student to get an idea of the magnitude of the sines of angles. For accurate computations a more extended table should be used.

Angle	0°	10°	20°	30°	40°	50°	60°	70°	80°	90°
Sine	0.000	.174	.342	.500	.643	.766	.866	.940	.985	1.00

591. Snell's law of refraction. In Fig. 386 the angle i_1 is the *angle of incidence*, and i_2 is the *angle of refraction*. The law which states the quantitative relation between these two angles is called *Snell's law*, or sometimes the *sine law of refraction*. This law can be stated by the following equation:

$$\frac{\sin i_1}{\sin i_2} = \text{a constant}. \qquad (10)$$

According to this law, if the angle of incidence i_1 is changed in size, the angle of refraction i_2 must change in such a way that the ratio of the sines of the two angles will remain constant.

The constant which appears in Snell's law is called the *index of refraction*. As will be shown in the next section this index has a simple physical meaning, but for the present it may be regarded as a mere number. This index is different for different media. In the case of light going from air into water it is approximately $\frac{4}{3}$.

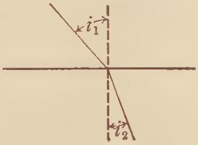


FIG. 386

To illustrate the use of this law a few cases will be solved for light traveling from air into water. Since the index is $\frac{4}{3}$,

$$\frac{\sin i_1}{\sin i_2} = \frac{4}{3},$$

hence $\sin i_1 = \frac{4}{3} \sin i_2$,

or $\sin i_2 = \frac{3}{4} \sin i_1$.

When $i_1 = 10°$,

$\sin i_1 = 0.174$

and $\sin i_2 = \frac{3}{4}(0.174) = 0.130$;

hence $i_2 = 7.5°$.

When $i_1 = 60°$,

$\sin i_1 = 0.866$

and $\sin i_2 = \frac{3}{4}(0.866) = 0.649$;

hence $i_2 = 40.5°$.

In each of these cases the ratio of the sines of the angles is constant; that is,

$$\frac{\sin 10°}{\sin 7.5°} = \frac{\sin 60°}{\sin 40.5°} = \frac{4}{3} \text{ (approximately)}.$$

592. Proof of the sine law by the wave theory. The wave theory explains refraction not only in a *qualitative* way but also in a *quantitative* way. The latter will be shown by deriving theoretically the quantitative law of refraction.

In Fig. 387 line AB represents a wave traveling in the direction of the arrow. After passing into a different medium, the wave changes to the position DC. While the part of the wave near B travels in the air to C, that part near A travels to D, a distance which is shorter on account of the slower rate at which light travels in the second medium. The line AD is drawn perpendicular to the wave DC, and hence represents the direction that light is traveling in the second medium.

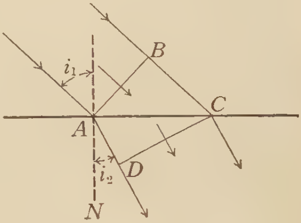


FIG. 387

In the figure the angle of incidence, i_1, is marked. Since the lines AB and AC are perpendicular, respectively, to the two sides of the angle i_1, the angle BAC is equal to i_1. For a similar reason, angle ACD is equal to i_2. In the triangle ABC, from the definition of the sine of an angle,

$$\sin BAC = \sin i_1 = \frac{BC}{AC}. \tag{11}$$

Similarly, in the triangle ACD,

$$\sin ACD = \sin i_2 = \frac{AD}{AC}. \tag{12}$$

Dividing equation (11) by equation (12), AC cancels out, giving

$$\frac{\sin i_1}{\sin i_2} = \frac{BC}{AD}. \tag{13}$$

But BC is the distance the light travels in the first medium while it is traveling the distance AD in the second medium; hence

$$\frac{BC}{AD} = \frac{\text{speed of light in first medium}}{\text{speed of light in second medium}}.$$

If we place V_1 for the speed of light in the first medium, and V_2 for the speed in the second,

$$\frac{BC}{AD} = \frac{V_1}{V_2},$$

and equation (13) becomes

$$\frac{\sin i_1}{\sin i_2} = \frac{V_1}{V_2}. \qquad (14)$$

The ratio V_1/V_2 is independent of the direction the light travels, and therefore remains constant as i_1 and i_2 are changed. Hence we have proved that the ratio of the sines is a constant. This is Snell's law of refraction.

We have not only deduced Snell's law (the sine law) but have found a physical meaning for the constant (index of refraction) of this law. *The index of refraction is equal to the ratio of the speeds of light in the two media.*

Since the index of refraction from air to water is approximately four thirds, and since the speed of light in air is about 300,000,000 meters per second, the speed of light in water is three fourths of 300,000,000, or 225,000,000, meters per second (approximately). The speed of light can be computed for any other medium when the index has been found by experiment. The observed speeds, for all cases where accurate measurements have been made, agree with the values found by computation.

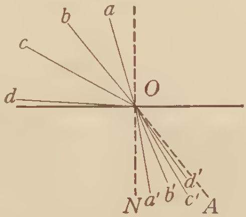


Fig. 388

593. Total reflection; the critical angle. Consider the case of light falling at different angles of incidence on the surface of a medium in which light travels more slowly than in the first medium. When the direction of light is that indicated by a ray a (Fig. 388), it will be refracted to the direction shown by a'; when it has the direction of the ray b, it will be refracted into b'; and so on. The ray d, which is almost tangential to the surface, becomes d' after refraction. *Under no circumstances does light pass into the second medium with an angle of refraction greater than AON.*

If a large obstacle, such as a raft, is floating on clear, deep water, no light from above will be refracted into the region lying to the right of and above the line *OA* of Fig. 389.

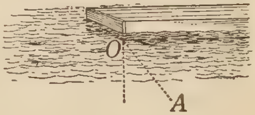


FIG. 389

In Fig. 390 is shown the range of directions from which light traveling through the air could enter an eye that is placed under water. The entire landscape will be seen in a direction inclosed by the angle *AEB*. Distant objects which are near the level of the surface of the water will apparently be elevated about 40 degrees.

Returning again to the case of Fig. 388, consider what would happen if the path of light were reversed. When light traveled in the direction *a′O*, it would be refracted to *a*; if it traveled in the direction *b′O*, to *b*; if in the direction *c′O*, to *c*; and if in the direction *d′O*, it would be refracted to *d*, almost tangentially to the surface. Theoretically, if the light traveled in the direction *AO*, it would be refracted tangentially to the surface. What would happen when the angle of incidence became greater than the angle *AON*?

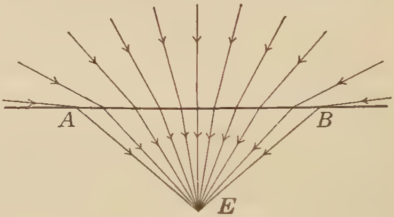


FIG. 390

Obviously the light would not emerge above the surface; for if it did, it should be possible to make it return by the same path, and we have seen that no light will return into the lower medium at an angle greater than the angle *AON*. Experiment shows that the light is *totally reflected*; that is, light which is traveling in the direction *BO* (Fig. 391) is reflected in the direction *OB′*, *none* of it passing through the surface.

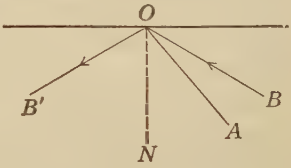


FIG. 391

Light is totally reflected whenever the angle of incidence exceeds a certain angle, called the critical angle, and when the incident light is in the medium in which light has the smaller velocity. For example, total reflection in the case of air and water occurs only when light attempts to go from water into air and when the angle of incidence is greater than about 49 degrees.

When light is totally reflected from a surface to the eye, the surface looks like a silvered mirror. For example, if one holds a tumbler of water at about the height of the top of his head and looks through the side of the glass at the underside of the surface of the water, it will look like a silvered surface. In that case it is totally reflected light that comes to the eye.*

A more rigorous explanation of total reflection based on the sine law of refraction will now be given. In the case of air and water

$$\frac{\sin i_1}{\sin i_2}=\frac{4}{3},$$

where the angle i_1 is the angle in air, and i_2 is the angle in water. Whether the angle i_1 is the angle of incidence or the angle of refraction depends on which way the light is traveling: if the light is incident in the water, traveling out, i_1 is the angle of refraction, and i_2 the angle of incidence. The sine law may be written

$$\sin i_1=\tfrac{4}{3}\sin i_2. \tag{15}$$

Suppose that light is incident inside the water at a known angle; that is, i_2 is the angle of incidence and is known. To find the value of sin i_1 one must look up the sine of i_2 and multiply it by four thirds. This works all right provided i_2 is less than about 49 degrees. Since the sine of 49 degrees is three fourths (approximately),

$$\sin i_1=\tfrac{4}{3}\times\tfrac{3}{4}=1.$$

When the sine of an angle is equal to 1, the angle is equal to 90 degrees. Hence, if light is incident in the water at an angle of 49 degrees, it will be refracted at an angle of 90 degrees, or tangentially to the surface. But when the angle of incidence is greater than 49 degrees, sin i_2 will be greater than three fourths, and sin i_1, as computed by equation (15), will be greater than 1. But no angle can have a sine greater than unity; hence there can be no angle of refraction, there cannot be refraction, and the light must be totally reflected. For an air-water surface the angle of 49 degrees is the *critical* angle.

In general, if the index of refraction is denoted by n, then

$$\sin i_1=n\sin i_2. \tag{16}$$

---

*Total reflection can easily be shown by the aid of a glass tank (Fig. 379).

When the angle i_1 is equal to 90 degrees, i_2 is equal to the critical angle. For this case ($\sin i_1 = 1$), equation (16) becomes

$$1 = n \sin i_2.$$

Hence

$$\text{Sine of the critical angle} = \frac{1}{n}. \tag{17}$$

CRITICAL ANGLES, WITH AIR AS ONE MEDIUM

	Index	Critical Angle
Carbon bisulfide	1.6	39°
Crown glass	1.5	42°
Diamond	2.5	24°
Flint glass	1.7	36°
Water	1.33	49°

594. Examples of total reflection. In the case of an air-water surface, total reflection takes place only when the incident light is inside the water. This is the case in the luminous jets sometimes seen in fountains. Fig. 392 shows one method of producing a luminous jet. Light enters through a glass window into a chamber through which water is flowing. As shown in the figure the light entering the water jet is totally reflected on the inside of the jet and follows it around a curve. There is enough light scattered or diffusely reflected in all directions from particles in the water to make the jet appear luminous.

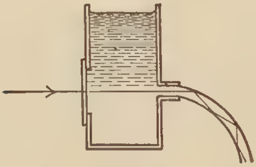


Fig. 392

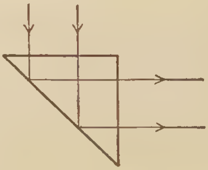


Fig. 393

The *totally reflecting prism* is used in many kinds of optical instruments, because it gives reflections superior to those from a silvered mirror. As shown in Fig. 393, a section of such a prism is an isosceles right-angled triangle. Light falling perpendicularly to one face enters the prism without any bending of the beam and strikes the inclined face at an angle of incidence of 45 degrees. Since this angle is greater than the critical angle of glass, the light is totally reflected, and passes out through the third face, as shown in the figure.

595. The relative index and the absolute index. The index of refraction between two media, or the *relative index*, is defined by the following equation:

$$\text{Relative index of two media} = \frac{\text{speed in first medium}}{\text{speed in second medium}}.$$

Hence the sine law of refraction may be stated thus:

$$\frac{\sin i_1}{\sin i_2} = \frac{\text{speed in first}}{\text{speed in second}} = \text{relative index},$$

where i_1 is the angle the ray makes with the normal to the surface in the first medium, and i_2 is the corresponding angle in the second medium.

The *absolute* index of a substance is defined as follows:

$$\text{Absolute index} = \frac{\text{speed in vacuum}}{\text{speed in substance}}.$$

Accurate values given in tables for the indexes of substances are the absolute ones. However, as the speed of light in air differs so little from that in a vacuum, it is only in very accurate work that one needs to distinguish between them.

In designing high-grade lenses it is necessary to know how much the light is bent in passing from one kind of glass to another. In that case one must know the *relative* index of refraction. But the tables give only the *absolute* indexes of different substances. How can one use the sine law and make the necessary computations? In answering this question not only will a method of computing the relative index be given but a more general form of the sine law will be stated. To make the discussion more concrete we shall take the case where light is traveling from water to glass.

The *absolute* index of water is, by definition,

$$n_1 = \frac{V_0}{V_1}, \tag{18}$$

where V_0 is the speed of light in a vacuum, and V_1 is the velocity in the water. The *absolute* index of glass is

$$n_2 = \frac{V_0}{V_2}, \tag{19}$$

where V_2 is the speed in glass. If equation (19) is divided by equation (18), we have

$$\frac{n_2}{n_1}=\frac{V_1}{V_2}. \qquad (20)$$

But the ratio V_1/V_2 is the relative index for the two substances designated as 1 and 2; hence the relative index is equal to the *inverse* ratio of the absolute indexes of the two substances.

The sine law of refraction can now be written in terms of the absolute indexes. From equations (14) and (20) we have

$$\frac{\sin i_1}{\sin i_2}=\frac{V_1}{V_2}=\frac{n_2}{n_1},$$

where the *n's* are the absolute indexes. This may be written

$$n_1 \sin i_1 = n_2 \sin i_2, \qquad (21)$$

a form very easy to remember on account of the symmetry of the subscripts.

Equation (21) may be called the *general law of refraction.* It is the equation that is used in computing the change in direction of light which is traveling from water to glass or from one kind of glass to another.

596. The prism. The prism is used so much in optics that the student should be certain that he understands why a ray follows the path shown in Fig. 394. The incident ray is bent toward the normal when it enters the prism and is bent away from the normal as it leaves. The result of this is that the angle of deviation (D in the figure) is large.

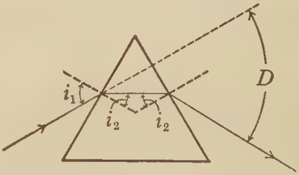


FIG. 394

597. Dispersion. Whenever a narrow beam of light from any ordinary source passes through a prism and then falls on a screen, a colored band, or *spectrum,* is produced (Fig. 395). To understand this three points must be clear:

1. It was proved by Sir Isaac Newton that sunlight is a mixture of different kinds of light. As will be explained later, these kinds differ in the fact that their wave-lengths, and hence their frequencies, are not the same.

2. Different kinds of light produce different color sensations in the eye. Color depends on the frequency of the light-vibrations. The lowest frequencies, or longest waves, of visible light produce

the sensation of red; while the highest frequencies, or shortest waves, produce the violet sensation. The principal colors in the spectrum, in the order of their frequencies, are red, yellow, green, blue, and violet.

3. The velocity of light in glass depends on the frequency of vibration of the waves. Red light travels faster in glass than those kinds which have a higher frequency. Violet light travels slower than blue, blue slower than green, green slower than yellow, and yellow slower than red.

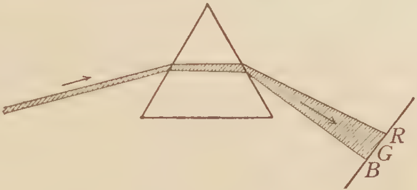


Fig. 395

Since the amount that light is bent by glass depends on the velocity in the glass, and since the velocity is different for different frequencies, it follows that when a beam of white light, which is a mixture of different frequencies, passes through a glass prism, the beam will be spread out as indicated in Fig. 395. Since the different frequencies produce different color sensations, the band formed on the screen is colored.

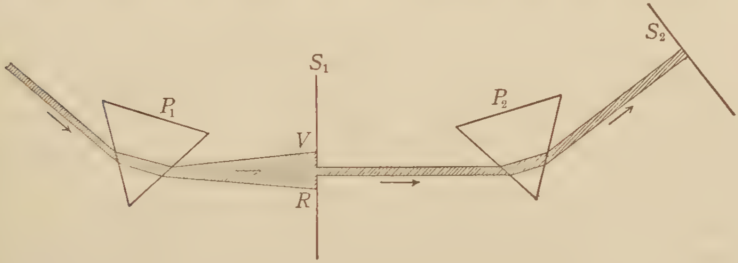


Fig. 396

The breaking up of a beam of light into its component parts is called *dispersion.*

598. Newton's experiment with colors. As stated in the last section, it was Newton who showed that sunlight is a mixture of different kinds of light. The method of his experiment was relatively simple. A narrow beam of sunlight, after passing through the prism P_1 (Fig. 396), was broken up into a spectrum on the screen

S_1. A small opening in this screen allowed the light of one color to pass through and fall on a second prism, P_2. After passing through this prism, it fell on the screen S_2. Newton found that the light which passed through the second prism P_2 was not separated into different colors. For example, if green light passed through the hole in the first screen, only green light was seen on the second screen; if the first screen was moved so that some other color, say red, passed through the hole, only red would be seen on the second screen. In this way Newton showed that the first prism resolved white light into parts which could not be further separated by a second prism. In other words, a prism can separate light into components each of which is homogeneous (not a mixture).

599. Atmospheric refraction of light. The velocity of light in air is slightly less than that in free space. The denser the air the less is the velocity. Light which reaches us from celestial objects travels down through an atmosphere of increasing density, and its velocity gradually changes as it approaches. When the light comes from directly overhead, there will be no refraction, or bending; but when the light comes from objects in other portions of the sky, there will be a change in its direction. As shown in Fig. 397, the bending is in such a direction that the object is apparently displaced toward the zenith; light coming from the direction A is bent so that it appears to come from A', and that from B appears to come from B'. The bending is very small except for objects located near the horizon. At an angular elevation of 45 degrees the displacement is about one thirtieth of the diameter of the moon; at an angular elevation of 15 degrees it is about one sixth the diameter of the moon. This increase of bending as objects get nearer the horizon is the cause of the apparent flattening of the sun and the moon at the horizon. The lower edge is raised more than the top, which makes the vertical diameter appear less than the horizontal one. The refraction at the horizon is enough to make an appreciable difference in the apparent time of the rising and the setting of the sun or moon. Inasmuch as the moon or the sun at the horizon is apparently raised by refraction by an amount a little more than its diameter, we can see the moon or the sun when it is really below the horizon.

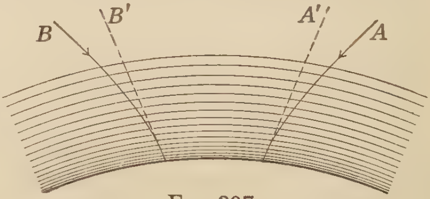


FIG. 397

When one looks over a hot radiator or a hot pavement, there is seen a wavy, shimmering effect. The density of hot air is less than that of cold air, and light travels faster in it when it is hot than when it is cold. For this reason, when a beam of light passes from cold air to hot air or from hot air to cold air, it is bent. Thus a ray of light passing over a hot radiator is bent in passing through the heated air. If this heated air would remain quiet, the amount that the light is bent would remain constant. But the hot air does not do this,—it is in constant motion. Moreover, it is being mixed with colder air, with the result that its density is continually changing. For this reason the amount that the ray is bent is varying continuously. Hence the apparent position of an object seen through this heated air is changing back and forth or up and down in a very irregular way. It is commonly said that one sees the "heat waves" rising. That is not so. The effect is caused by the change in the amount that light is bent as it travels through air of different densities. The same thing can be seen in looking through a glass vessel of water into which glycerin or alcohol is being poured.

The *twinkling*, or *scintillation*, of stars is due to the unstable condition of the air through which the light comes, and hence to the change in the amount that the light is bent. Since there is a slight difference in the velocity of different colors, these may be bent differently, and there may be changes in the color of the star as it twinkles. Light from distant terrestrial sources of light will often twinkle when there is a mixing of warm and cool air currents between the source and the observer.

When currents of warm and cold air are mixing on a large scale, there is considerable interference with good seeing. This is of very great importance in telescopic work. It is sometimes called *optical haze* and distinguished from *dust haze*. It may be very noticeable on hot days, even when the air is comparatively free from dust.

A number of rather curious distortions of distant objects near the horizon occur under special conditions. Among these distortions are the phenomena known as *looming*, *towering*, *sinking*, and *stooping*.* They can all be explained by refraction.

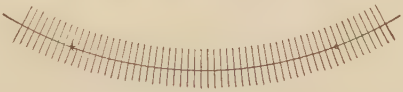


FIG. 398

600. The mirage. A number of different optical effects, known under the general name of *mirage*, are due to atmospheric refraction. Only one kind, known as the *inferior mirage*, will be explained here—the kind most commonly observed.

Under special circumstances, as in the case of a heated pavement or a dry plain, the density of the air is less near the surface. In that case light-waves are refracted as shown in an exaggerated way in Fig. 398. In Fig. 399

---

* See Humphrey's "Physics of the Air."

the eye is at E. Light from the sky coming down over the tree tops enters the hot air near the heated pavement and is bent up to the eye. Some light from the trees will travel straight across through the cooler air. The result is that some of the sky will be seen apparently *below* the trees. Sky seen in that direction will appear as sky reflected in water. In some cases the light from the trees which travels down toward the pavement may be bent up, and thus one will apparently see images of trees in the water. On a bright, clear day it is usually possible to find some street or pavement which, when one stoops down and looks along it, will show some sky light *below* the distant trees or other objects. It is a common experience of motorists, on clear days, for a paved roadway some distance in front of them to appear covered with water.

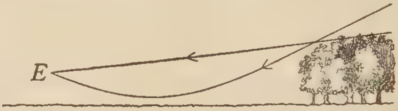


FIG. 399

PROBLEMS

1. The angle of incidence on glass is 50°, and the observed angle of refraction is 35°. Find the index of refraction. ($\sin 50° = 0.766$; $\sin 35° = 0.574$.)

2. The angle of incidence of light on water, index 1.33, is 30°. Find the sine of the angle of refraction.

3. Light which emerges from water, index 1.33, has an angle of incidence in the water of 40°. Find the sine of the angle of refraction.

4. The index of refraction of a certain kind of glass is 1.5. Compute the speed of light in this glass.

5. For light which has a wave-length in air of 0.00006 cm., what would be the wave-length in glass the index of which is 1.5?

6. A block of glass the index of which is 1.5 is completely submerged in water (index 1.33). Light strikes the glass at an angle of incidence of 30°. Compute the sine of the angle of refraction in the glass.

7. A 60°-prism of crown glass, index 1.5, is submerged in carbon bisulfide, the index of which is 1.6. Make a diagram showing the approximate path of a ray of monochromatic light through the prism.

8. Draw a diagram similar to that requested in Problem 7, but for the case of the prism submerged in water, the index of which is 1.33.

CHAPTER XLIII

LENSES

Converging and diverging lenses, 601. The principal axis of a lens, 602. The principal focus, 603. The optical center, 604. Graphical construction of images, 605. The quantitative law of position of image, 606. Derivation of the quantitative law, 607. The quantitative law applied to diverging lenses, 608. Summary of rules for use of signs in the quantitative law, 609. Law of size of image, 610. Two thin lenses in contact, 611. The power of a lens; the dioptric, 612. Spherical aberration, 613. Chromatic aberration, 614. Astigmatism, 615. Distortion of images, 616.

601. Converging and diverging lenses. When a beam of parallel light passes through a lens having convex surfaces (Fig. 400), the beam is converged to a focus. A lens of this type is said to be a *converging* lens. But when a beam of parallel light passes through a lens having concave faces (Fig. 401), the emerging beam diverges, and the focus is a virtual one. Such a lens is called a *diverging* lens.

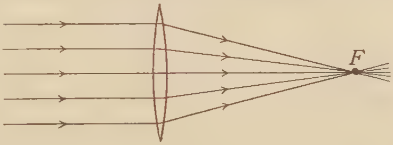


FIG. 400

It is very important to understand the cause of these effects. The converging or diverging power of lenses depends on the fact that light travels more slowly in glass than in air. In the beam incident on the lens of Fig. 402 the waves are shown by the straight parallel lines. Because light travels more slowly in glass than in air, that part of the waves which travels through the thicker portions of the lens is retarded more than the part which travels through the thinner portions;

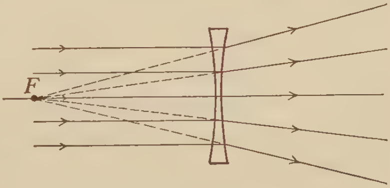


FIG. 401

hence the emerging waves are curved as shown in the figure. For thin lenses with spherical surfaces the emerging waves will be spherical, except in cases explained later (sect. 615). Since light always travels perpendicularly to the waves, it converges to their center of curvature (*F* in the figure).

In a diverging lens the glass is thinner at the center and thicker near the edges; hence the central portion of each wave gets ahead (Fig. 403). Since light travels perpendicularly to the waves, the beam diverges from their center of curvature (*F* in the figure).

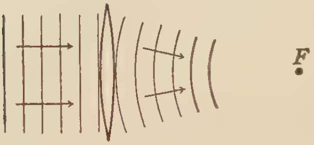


FIG. 402

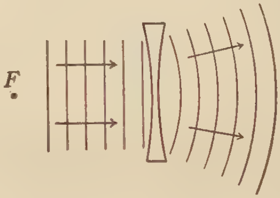


FIG. 403

From the foregoing explanations it should be clear that in a converging lens the glass is thicker at the center, and that in a diverging lens it is thinner at the center. This statement is a rule for telling which lenses are converging and which diverging. In Fig. 404 are examples of both types.

602. The principal axis of a lens. The principal axis of a lens is a line drawn through the center of the lens and the centers of curvature of the faces.

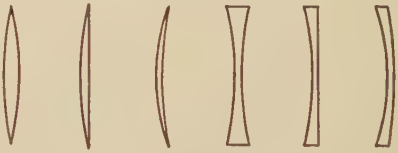


FIG. 404

603. The principal focus. The principal focus of a lens is defined in a manner similar to that used for mirrors. The *principal focus of a lens is the point which is the focus when the incident light is a beam parallel to the axis* (Figs. 400 and 401).

In a converging lens the principal focus is *real* (Fig. 400); in a diverging lens it is *virtual* (Fig. 401). If a *point* source of light could be placed at the principal focus of a converging lens, the light, after going through the lens, would form a parallel beam.

There are always two principal foci for a lens, one on each side; for light can travel in either direction through the lens.

604. The optical center. In a *thin* lens a certain point has the property that the rays which pass through it are not appreciably bent by the lens. This point is called the *optical center*. If the lens is a symmetrical one, such as a double-convex or double-concave lens, the optical center is at the geometrical center. However, the statement that a ray which passes through the optical center is not bent is an approximation, and is true only for thin lenses. If the lens is thick, there are points, called *nodal points* (M and N in Fig. 405), which have the following properties: Any incident ray which travels toward the nearer one of these points is so refracted that it emerges traveling parallel to its original direction, in a line which, if drawn back into the lens, would pass through the other nodal point. The path of such a ray is shown in the figure. In thin lenses the two points practically coincide in a point called the optical center. *In a thin lens all rays which pass through the optical center emerge from the lens without having been appreciably bent.*

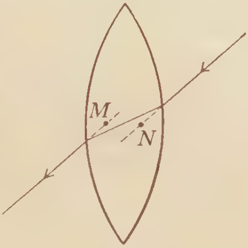


FIG. 405

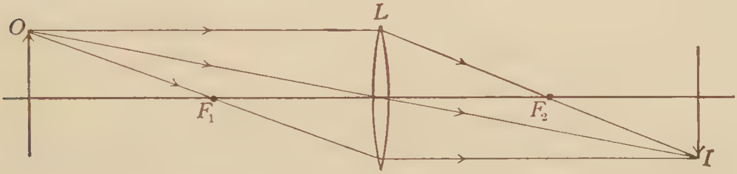


FIG. 406. Real image formed by a converging lens

605. Graphical construction of images. In most cases it is comparatively easy to locate graphically the position of an image formed by a thin lens. In Fig. 406, O is the object. From the upper end of the object are drawn three rays the paths of which are known. First, a ray which is parallel to the axis will, after refraction, go through the principal focus F_2 as shown. A second ray is drawn straight through the optical center of the lens. As stated in the last section, such a ray is not appreciably bent. A third ray passes through the principal focus F_1; because it goes through that point, it travels parallel to the axis after emerging from the lens. These three rays meet at a point forming the image of the point from which they started. The image is *real* and *inverted*. In actual practice it is not necessary to draw three rays: any two of the three are sufficient. The object and the image of Fig. 406 are at *conjugate foci*, and their positions may hence be interchanged.

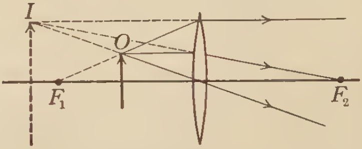


FIG. 407. Virtual image formed by a converging lens

In Fig. 407 is given the graphical construction for an object inside the principal focus. Here, as before, the image of only one point of the object is found. To attempt more would only cause confusion. Three rays are drawn as before. The ray parallel to the axis is bent down so that it passes through the principal focus F_2; the second is drawn straight through the optical center of the lens; and the third, which has the same direction as if it came from F_1, travels, after emergence, in a direction parallel to the axis. It is seen in the figure that the three emerging rays are divergent and will never meet. However, they appear to come from the virtual image *I*. This virtual image is *erect* (that is, right side up) and *magnified*.

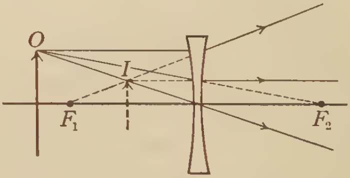


FIG. 408. Virtual image formed by a diverging lens

The two graphical constructions which have just been given show (1) that a *real* image is obtained when *the object is outside the principal focus*, and (2) that a *virtual* image is obtained when *the object is inside the principal focus*. It is easy to verify these statements by simple experiments with a lens.

Diverging lenses, except when used in conjunction with other lenses, never form real images. No matter where the object is, the image is always virtual. This may be proved by the graphical method (Fig. 408). In this type of lens the principal foci, F_1 and

F_2, are *virtual* and must be treated differently from those of Figs. 406 and 407. In this figure the ray drawn from O parallel to the axis is bent so as to *appear* to come from F_1; the ray drawn from O *toward* F_2 will be parallel, after emergence, to the axis. In this case the emerging rays are divergent, and appear to come from an image at I. The image is *virtual*, *erect*, and *diminished*.

606. The quantitative law of position of image. There is a law, similar to that for mirrors, which gives the *quantitative* relation between the distances of the object, the image, and the principal focus from the lens. If p is the distance of the object from the lens, q the distance of the image from the lens, and f the distance of the principal focus from the lens, then

$$\frac{1}{p}+\frac{1}{q}=\frac{1}{f}. \tag{22}$$

As a numerical illustration suppose that a lens which has a focal length (the distance from the lens to the principal focus) of 30 centimeters is placed 40 centimeters from the object. Where will the image be? In this case $p=40$ and $f=30$. Substituting these values in equation (22), we obtain

$$\frac{1}{40}+\frac{1}{q}=\frac{1}{30};$$

hence

$$\frac{1}{q}=\frac{1}{30}-\frac{1}{40}=\frac{1}{120}$$

and

$$q=120 \text{ cm.}$$

Hence the image will be located 120 centimeters from the lens. The two points—one 40 centimeters and the other 120 centimeters from the lens—are conjugate foci. If the object is placed 120 centimeters away, the image will be found at the 40-centimeter point. This can easily be proved by equation (22).

The student should verify the following propositions for the case of a lens with a focal length of 30 centimeters:

When $p=240$ cm.,
$q=34.3$ cm.
When $p=60$ cm.,
$q=60$ cm.
When $p=30$ cm.,
q is infinitely large.

In the last case the object is at one of the principal foci.

It has been pointed out that when the object is inside the principal focus, the image is virtual. In terms of the notation of equation (22), this means that p is less than f. In general, we have, from equation (22),

$$\frac{1}{q}=\frac{1}{f}-\frac{1}{p}.$$

From this relation it can be seen that if p is less than f and hence $1/p$ is larger than $1/f$, the right-hand side will always be negative. Hence, *in the use of equation (22) a negative value of q means a virtual image.*

607. Derivation of the quantitative law. The relation given by equation (22),

$$\frac{1}{p}+\frac{1}{q}=\frac{1}{f},$$

will now be derived. In Fig. 409, AB is the object and EG the image. The parallel ray AC passes, after refraction, through the principal focus F. The ray AD passes through the optical center of the lens without appreciable bending. By construction, $AB = CD$. The two triangles CDF and FEG are similar; hence

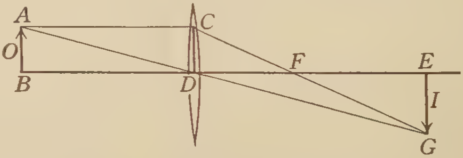


FIG. 409

$$\frac{DF}{FE}=\frac{CD}{EG}=\frac{AB}{EG}.$$

Since the triangles ADB and DEG are similar,

$$\frac{BD}{DE}=\frac{AB}{EG}.$$

The left-hand members of the last two equations are both equal to the same ratio; hence they must be equal to each other:

$$\frac{BD}{DE}=\frac{DF}{FE}. \qquad (23)$$

Expressed in terms of the adopted notation,

$$BD = p,$$

$$DE = q,$$

$$DF = f,$$

and

$$FE = q - f.$$

Substituting these values in equation (23),

$$\frac{p}{q} = \frac{f}{q - f}.$$

Clearing of fractions, $qf = pq - pf.$

Dividing this equation by pqf, we have

$$\frac{1}{p} = \frac{1}{f} - \frac{1}{q}.$$

From this, by a simple transposition, we have the desired equation (equation (22)),

$$\frac{1}{p} + \frac{1}{q} = \frac{1}{f}.$$

It should be noticed that in the construction of Fig. 409 it was assumed that the lens was *thin*. In general, the methods explained in this chapter apply only to thin lenses.

608. The quantitative law applied to diverging lenses. In a diverging lens the principal focus is a virtual one (see Fig. 401). In section 606 it was pointed out that when the image is virtual, the distance q is a *negative* quantity. For a similar reason the focal length f of a diverging lens is negative. Since the lenses are often rated in terms of their focal lengths, a diverging lens is often called a *negative* lens.

*The quantitative law given by equation (22) can be applied to diverging lenses if the focal length is taken as negative.**

---

* It is best to form the habit of always writing the quantitative law in the form in which it appears in equation (22). If the image or principal focus is virtual, the negative sign should be used when the numerical value is inserted in the equation; in other words, let the negative sign always go with the numerical quantity and not with the algebraic quantity. This simple rule will enable one to avoid much confusion.

The focal length of a diverging lens is 30 centimeters, and the object is 40 centimeters away. Where is the image? In this case $f = -30$ and $p = 40$.

$$\frac{1}{40} + \frac{1}{q} = -\frac{1}{30},$$

or

$$\frac{1}{q} = -\frac{1}{30} - \frac{1}{40} = -\frac{7}{120};$$

hence

$$q = -\tfrac{120}{7} = -17.1 \text{ cm.}$$

Hence the image is *virtual* and 17.1 centimeters from the lens.

It can readily be seen by looking over the numerical work just given that q will come out negative, no matter what the numerical value of p is, provided that p is a positive quantity. In other words, no matter where the object is, *a diverging lens, like a diverging mirror, when used alone always produces a virtual image.*

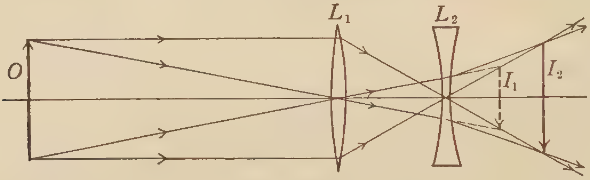


FIG. 410

There is one instance where a real image can be produced by a diverging lens; namely, where a converging lens is used with a diverging one. In Fig. 410 the lens L_1 would, if used alone, produce an image at I_1. But the lens L_2 diverges the light; hence the image is not formed at I_1 but at I_2. In this case a real image is formed by the diverging lens. The image I_1 is really the *object* for the lens L_2, and is a *virtual* object. Equation (22) will apply if p, the distance of the object from the lens, is taken as a *negative* quantity whenever the object is virtual, as in Fig. 410.

609. Summary of rules for use of signs in the quantitative law. 1. In a converging lens the focal length is positive (the principal focus is real); in a diverging lens the focal length is negative (the principal focus is virtual).

2. When the image is real, q is positive; when it is virtual, q is negative.

3. The quantity p is always positive except in those special cases where the object is a virtual one (see the last paragraph of section 608).

610. Law of size of image. By the aid of Fig. 411 it should be obvious that

$$\frac{\text{Size of object}}{\text{Size of image}} = \frac{p}{q}. \qquad (24)$$

This is the law of the relative size of object and image. The student should be able to show, from the figures already given, that this law applies to virtual images as well as to real images. It will also appear (in Fig. 407, for example) that the virtual image made by a converging lens is always larger than the object, while the virtual image made by a diverging lens (as in Fig. 408) is always smaller than the object.

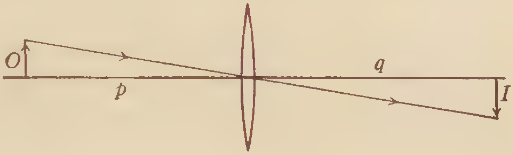


FIG. 411

611. Two thin lenses in contact. It can be proved that when two thin lenses are in contact,

$$\frac{1}{p} + \frac{1}{q} = \frac{1}{f_1} + \frac{1}{f_2},$$

where f_1 is the focal length of one lens, and f_2 the focal length of the other lens. We may regard this combination as equivalent to a single lens of focal length F. Then

$$\frac{1}{F} = \frac{1}{f_1} + \frac{1}{f_2}. \qquad (25)$$

612. The power of a lens; the dioptric. Instead of describing lenses in terms of the focal length, opticians use the reciprocal of the focal length and call this reciprocal the *power of the lens.*

When the focal length of a lens is measured in meters, the reciprocal of this gives the power measured in a unit called the *dioptric* (sometimes written *diopter*); in other words, the reciprocal of the focal length of a lens measured in meters is equal to the number of dioptrics. For example, a converging lens with a focal length of 0.5 meter is rated as + 2.0 dioptrics. A lens marked − 1.5 dioptrics has a focal length of 0.67 meter, and is a diverging lens.

There are several advantages in defining the power of a lens as the reciprocal of the focal length. The following example is given for the case of two thin lenses in contact:

If P_1 is the power of one lens, P_2 the power of the second, and P the power of the combination, it follows from equation (25) that

$$P = P_1 + P_2,$$

a formula much used by opticians.

613. Spherical aberration. The methods which have been used in this text in the discussion of the formation of images by lenses are not rigorous. For example, it has been *assumed* that light which goes through the edges of a lens comes to the same focus as light which goes through the center. This is not strictly true, for the light which goes through the edges of a converging lens is converged a little more than that which goes through the center. This is shown in an exaggerated way in Fig. 412. The result of this is that images formed by a *wide* lens are somewhat blurred. This effect is called *spherical aberration.*

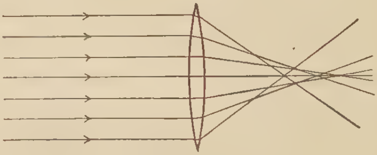


FIG. 412. Spherical aberration

It is relatively easy to show this aberration experimentally. If a lens which is relatively large in diameter is used to focus an image of some bright object on a screen, it will be seen that the image is much sharper when all the lens is covered except a small part of the center. The covering of part of the lens by a diaphragm is the usual remedy for spherical aberration. The stop, or diaphragm, in a photographic camera has as part of its purpose the correction of spherical aberration.

614. Chromatic aberration. In the last chapter it was shown that white light is not homogeneous, and that different components, producing different color effects, are refracted to different extents by glass and other refracting substances. Since violet light is bent more than red light, the focal length of a lens for violet light is less than that for red light, the difference being, for crown glass, about 4 or 5 per cent of the focal length. When ordinary light is used, the images formed by simple lenses have colored edges and sometimes so much general blurring that finer details of the image are lost. This form of aberration is called *chromatic aberration.*

To understand how lenses are corrected for chromatic aberration it is necessary to distinguish between two things: (1) the amount that the light is bent (that is, the *deviation* of the beam) and (2) the separation of the different colors (that is, the *dispersion*).

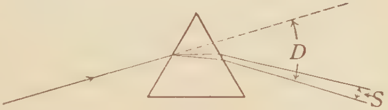


FIG. 413

These two effects are shown in Fig. 413. The angle D is the angular deviation of the beam, while the angle S shows the angular separation of the colors.

Different kinds of glass produce relatively different effects. Some kinds give a relatively large dispersion, others a relatively large deviation.

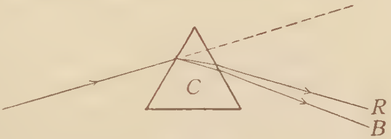


FIG. 414

The fact that different kinds of glass can be made which give the *same dispersion* but *different deviations* makes it possible to correct lenses for chromatic aberration. The principle can be understood by the aid of Figs. 414 and 415. The prism C, of crown glass, bends the beam and, moreover, produces a color separation as shown. The prism F, of flint glass, produces the *same* angular separation of the colors as prism C, but does not give so large a deviation of the beam. When these two prisms are combined, as in Fig. 416, *deviation is produced without dispersion.*

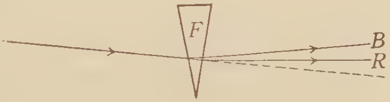


FIG. 415

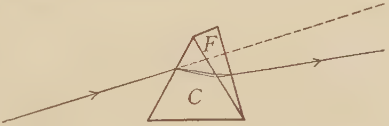


FIG. 416. Achromatic refraction

In Fig. 417 is shown an *achromatic combination* of two lenses, the converging lens being of crown glass, and the diverging lens of flint glass. If the converging lens alone had been used, the blue light from the source O would have converged to the point B, and the red light to the point R (in the figure the interval separating these foci is exaggerated); but when the diverging lens of flint glass is added, both the blue light and the red light are made to converge to the point *I*.

In practically all optical instruments achromatic combinations of lenses are used instead of simple lenses.

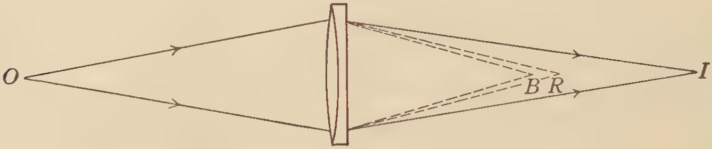


FIG. 417. An achromatic lens

615. Astigmatism. When light passes obliquely through an ordinary lens, horizontal and vertical lines in the object will not be brought to a focus in the same plane. A simple experiment can be arranged to show this. In Fig. 418 the object *O* is a lantern slide with horizontal and vertical lines on it, or a piece of "hardware cloth," which is woven wire with a coarse mesh. The object *O* should be illuminated in the way lantern slides usually are (the source is not shown in the figure). When an attempt is made to focus this image on the screen by a lens placed obliquely, it will easily be seen that the horizontal and vertical lines do not focus at the same place. One may be in focus at *A* and the other at *B*. Or, if a fixed position of the screen is used, the lens will have to be moved to different positions for focusing the two sets of lines.

This property which makes it impossible to focus horizontal and vertical lines at the same place is called *astigmatism*. In most optical instruments light does not travel obliquely enough through the lens to produce a noticeable astigmatism; but in the use of the photographic camera the light which forms those parts of a photograph that lie near the ends and corners must go obliquely through the lens. Hence for certain kinds of photography, astigmatism is a serious evil.

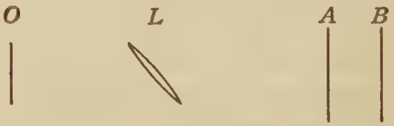


FIG. 418

A combination of lenses of different kinds of glass can be so assembled that almost complete correction for astigmatism is obtained. Such a combination is called an *anastigmatic* lens.

In spherical lenses astigmatism is shown when light travels obliquely through the lens; but it is produced also by a lens in which one or both surfaces are not spherical but curved in a manner similar to the side of a barrel, which has a greater curvature in one plane than in another. Often the human eye does not have spherical surfaces. In such cases the eye is not able to focus at the same time two sets of parallel lines if the two sets are at right angles to each other.

616. Distortion of images. The image of a rectangular mesh, such as a piece of "hardware cloth," may show either of the two types of distortion shown in Figs. 419 and 420. The first of these is sometimes called *barrel-shaped* distortion; the second, *pincushion-shaped* distortion. In the first case the center of the image is magnified more than the outer portions; in the second case the magnification is greater in the outer portions.

Distortion such as is shown in these figures can be controlled by the use of a properly placed diaphragm, or stop. In the photographic camera, if the diaphragm is placed in front (outside) of the lens, barrel-shaped distortion is produced; but if it is placed inside the lens (between the lens and the photographic film), pincushion-shaped distortion is produced. When the diaphragm is very close to the lens, no appreciable distortion is produced. A photographic lens which has its diaphragm so placed that there will be no distortion is called a *rectilinear* lens.

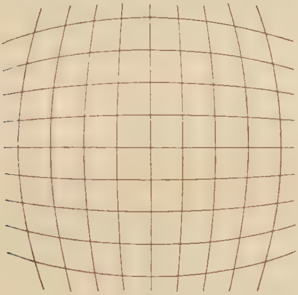


FIG. 419

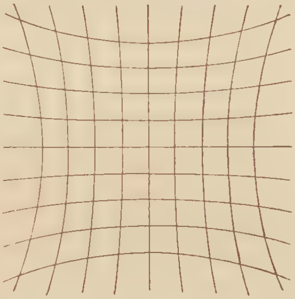


FIG. 420

The type of distortion shown in Figs. 419 and 420 is usually prevented by placing a diaphragm close to the lens; if a combination of lenses is used, distortion is prevented by placing the diaphragm inside the combination.

PROBLEMS

1. A converging lens has a focal length of 30 cm. (*a*) Where will the image be when the object is 60 cm. from the lens? (*b*) 40 cm. from the lens? (*c*) 20 cm. from the lens? (*d*) Which of these are real images?

2. A converging lens of focal length 20 cm. has a virtual image 20 cm. from the lens. Compute the position of the object.

3. Solve Problem 2 by a graphical method.

4. An object placed 40 cm. from a converging lens has a real image 60 cm. from the lens. Compute the focal length.

5. An object placed 20 cm. from a converging lens has a virtual image 40 cm. from the lens. Compute the focal length.

6. A diverging lens has a focal length of − 20 cm. (*a*) Where will the image be when the object is 40 cm. from the lens? (*b*) 20 cm. from the lens?

7. Solve Problem 6 by a graphical method.

8. An object is 120 cm. from a screen and 20 cm. from a lens. What is the focal length of the lens if a real image is formed on the screen?

9. A lens having a focal length of 10 in. projects an image of a lantern slide on a screen 30 ft. from the lens. If the slide is 3 in. wide, how wide is the image?

10. A lamp and a screen are 120 cm. apart. What must be the focal length of a lens that will produce a real image of the lamp twice as large as the source?

11. A diverging lens placed 10 cm. from an object produces an image which is half the size of the object. Calculate the focal length of the lens.

12. An object is 60 cm. from a converging lens which has a focal length of 30 cm. A second converging lens, with a focal length of 10 cm., is placed 68 cm. beyond the first lens. Find the position of the final image.

13. A converging lens with a focal length of 40 cm. is placed 35 cm. from a diverging lens with a focal length of − 10 cm. Compute the position of the image if the light from a far-distant object enters the converging lens first.

14. An optician places two thin lenses in contact. If one is + 1.25 dioptrics, and the other + 1.5 dioptrics, (*a*) what is the power of the combination? (*b*) what is the focal length of the combination?

15. (*a*) What is the power in dioptrics when a thin lens of − 1.25 dioptrics is placed in contact with a lens of + 2.00 dioptrics? (*b*) What is the focal length of the combination?

16. If one has a lens of + 4.0 dioptrics and wishes to form a combination which will have a focal length of 1 m., what power should the second lens have?

17. An achromatic lens is made by combining a thin lens which has a focal length of + 12 cm. with another thin lens which has a focal length of − 24 cm. What is the focal length of the combination?

CHAPTER XLIV

OPTICAL INSTRUMENTS

Introduction, 617. The photographic camera, 618. The eye, 619. The projection lantern, 620. The magnifying glass, 621. The compound microscope, 622. The astronomical telescope, 623. The terrestrial telescope, 624. The opera glass, 625. The prism binocular, 626. The reflecting telescope, 627. The range-finder, 628.

617. Introduction. It was shown in section 614 that mounting two lenses of certain kinds of glass together produces a combination which is, at least partially, corrected for chromatic aberration. It is shown in advanced treatises on lenses that all the different kinds of aberration can be approximately corrected by proper combinations of lenses. Hence, in most optical instruments, what is usually spoken of as a lens is in reality a group of lenses mounted together. Such a combination is equivalent to a single lens as far as the position and size of the image are concerned. It is best in elementary discussions of optical instruments to replace complicated combinations by simple lenses. Hence, in this text, instruments are usually simplified for purposes of explanation.

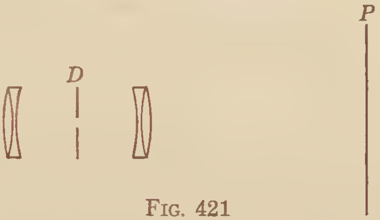


FIG. 421

618. The photographic camera. In the photographic camera a converging lens forms a real, inverted image of the object on a sensitive plate, or film (P of Fig. 421), which is inclosed in a light-tight case. The figure shows a common combination of lenses for a camera. There are two pairs of lenses, and each pair is corrected for chromatic aberration. Between these, at D, is the diaphragm, or the stop, as it is often called. Such a combination of lenses forms a system which for convenience may be regarded as a single lens.

The student should recall that images of objects which are at a great distance from a lens are focused at or near the principal focus, and that the images of near objects are focused some distance beyond the principal focus. Hence, in photographing distant objects the distance between the lens and the sensitized film is so adjusted that the film is at the principal focus of the lens system, but in photographing objects near at hand the lens must be moved farther from the film.

In the use of a camera having a short-focus lens, nearly all objects are relatively far away as compared with the focal length of the camera, and are imaged at or near the principal focus. In this case the camera may be focused once for all, giving a "fixed-focus camera," a common type in the cheaper grades. But in general the distance between the plate and the lens must be adjusted to suit the distance of the object which is to be photographed.

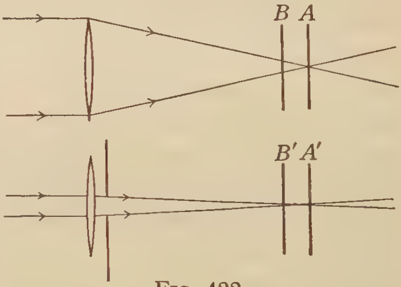


Fig. 422

The stop in a camera lens serves several useful purposes. When it is placed at the center of the lens system, as shown in Fig. 421, the distortion referred to in section 616 is eliminated, and all straight lines in the object are imaged as straight lines. The lens is then a *rectilinear* one. When a relatively small-sized stop is used, the image becomes sharper. Not only is the spherical aberration decreased (sect. 613) but other types of aberration are at least partly remedied. A small stop is also of great assistance in bringing both near and distant objects into focus at the same time. When the stop of a lens is opened wide, only certain objects are in focus on the film for any one position of the lens. For example, when an object thirty feet away is in focus, objects nearer and farther away are out of focus, being often rather blurred, and the lens is said to have only a small *depth of focus*. When a small stop is used, objects both far and near become much more distinctly imaged, and the depth of focus is improved. This effect can be understood by the aid of Fig. 422. When the lens is wide open, the light from some object converging to some one point of the image in the plane *A* forms a large angle. In the plane *B* this light is spread over a considerable area; and if the plate were at *B*, the image would be blurred.

When a small stop is used, the beams of light form small angles, and the light forming a point of an image in the plane A' will be spread over only a very small area if the plate is at B'. Hence the blurring of the image will not be so great in the position B' as in B. However, a small stop decreases the amount of light that reaches the photographic plate, or film, and hence a longer time of exposure is required.

619. The eye. The eye is in principle somewhat similar to the camera. A lens with its diaphragm, the *iris*, forms a real image on the *retina*, which is the screen at the rear. In the camera the focusing is done by moving the lens nearer or farther away from the plate or film; but in the eye the distance between the lens and the retina either is fixed or changes very little. The adjustment, or focus, in the eye is produced chiefly by changing the focal length of the lens by altering its curvature. This power of changing the shape of the lens of the eye is called *accommodation*. Older persons lose to some extent this power of accommodation, and need to wear glasses for reading, although they may see objects at a distance clearly and distinctly. What type of lens do they need for their glasses?

A shortsighted person can focus only on near objects, because the eye lens has too short a focal length, or because the distance between the lens and the retina is too great. By wearing diverging lenses the effective focal length of the combination of eye lens and glass lens is increased, and one can then see distinctly objects at a greater distance. When a person is farsighted, the focal length of the eye lens is too great, or the distance between the lens and the retina is too small, and he must wear converging lenses (the student should be certain that he understands why this type is necessary).

620. The projection lantern. Projection lanterns are now in common use. Not only are they used in classrooms and lecture rooms for projecting lantern slides but their use in connection with motion pictures has made them very popular. They contain two groups of lenses: L_1, the condenser lenses, and L_2, the projection-lens combination (Fig. 423). The projection lens L_2 consists usually of two pairs of lenses or some other combination of lenses giving a system corrected for several types of aberration. The lens L_2 is used to form a real image of the picture, S, on the screen. The screen is not shown in the figure, but must be at a considerable distance to the right. When lantern slides or the films of motion pictures are used, they are illuminated by transmitted light; that is to say, by the light which comes through them. In the figure the light from the source, O, is converged by the *condenser* lenses L_1 through the lantern slide, S, to the lens L_2. The use of condenser lenses greatly increases the amount of light which goes through the lens L_2 to the screen, and hence brightens the image.

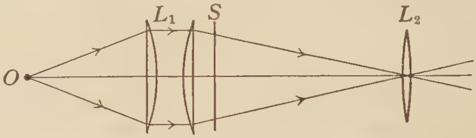


FIG. 423. The projection lantern

The focusing of the image on the screen is adjusted by moving the lens L_2 closer to or farther away from the slide S. The adjustment of the beam of light so that it converges toward the center of the lens L_2 is done by moving the source O.

621. The magnifying glass. In section 605 it was shown that when the object lies inside the principal focus of a converging lens, a virtual, erect, and magnified image is produced. The eye, looking through the lens (Fig. 424), sees at I an enlarged virtual image of O. Hence a simple lens may be used as a simple microscope, or magnifying glass.

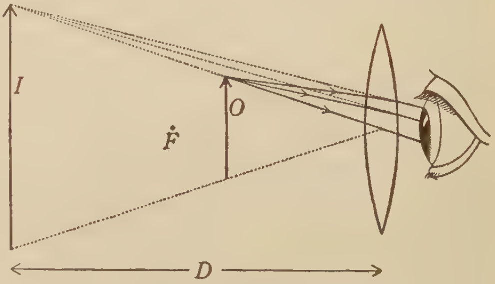


FIG. 424. The magnifying glass

There is a certain distance which is called the *distance of most distinct vision*. For the normal eye this is about 10 inches, or 25 centimeters. Hence, in using a magnifying glass, the best adjustment is made when the distance between the object and the lens is changed until the image I is at the distance of most distinct vision. It is important that the lens should be held close to the eye; for not only is the field of view larger, but the light entering the eye passes through the center of the lens, reducing the distortion and the spherical and chromatic aberration.

The *magnifying power*, m, of a simple microscope is defined by the expression

$$m = \frac{\text{size of image}}{\text{size of object}}.$$

From the law of the relative size of object and image (sect. 610), m is also equal to the ratio of the distances of the image and the object from the lens; that is, to the ratio q/p.

The law of lenses (equation (22)) may be written as

$$\frac{1}{p} - \frac{1}{q} = \frac{1}{f}$$

when it is remembered that in this case the image is virtual, and hence that q is negative. Multiplying this equation by q, we have

$$\frac{q}{p} - 1 = \frac{q}{f},$$

or

$$\frac{q}{p} = \frac{q}{f} + 1.$$

Hence the magnifying power, m, of a magnifying glass is

$$m = \frac{q}{p} = \frac{q}{f} + 1.$$

But if the eye is held close to the lens, and the optimum adjustment of focus is made, q becomes equal to D, the distance of most distinct vision.

Hence

$$m = \frac{D}{f} + 1, \tag{26}$$

where D is the distance of most distinct vision, and f the focal length of the lens used.

In the case of a lens the focal length of which is 2 centimeters, the magnifying power is

$$m = \tfrac{25}{2} + 1 = 13.5,$$

where D has been assumed to be 25 centimeters.

Equation (26) shows that the shorter the focal length of the lens, the greater the magnification. In order to improve the definition of the image, it is customary to use a group of long-focus lenses instead of a single lens with a short focus. Such a group has a focal length equivalent to that of a single short-focus lens.

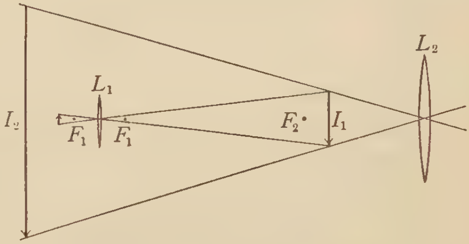


FIG. 425. The compound microscope

622. The compound microscope. Two groups of lenses are contained in the compound microscope. These groups are represented in Fig. 425* by the simple lenses L_1 and L_2. The object is placed just outside the principal focus of the *objective* lens, L_1, which has a short focal length. The lens L_1 produces a real image at I_1, which is formed inside the principal focus of the eyepiece lens, L_2. The eye, looking through the lens L_2, sees a magnified image, I_2, of the image I_1. The eyepiece lens is thus used as a magnifying glass to view the real image I_1.

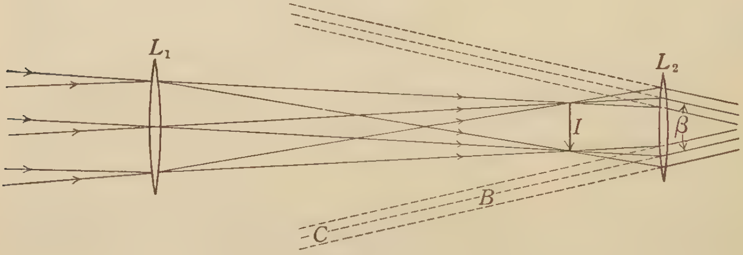


FIG. 426. The astronomical telescope

623. The astronomical telescope. In the astronomical telescope the objective is a long-focus lens (L_1, Fig. 426). Since the object viewed is at a great distance, the image I is at, or very near, the

---

[*] The student should recognize the method of construction used in this figure, for he need only recall that any line drawn through the optical center of a lens from a point on the object will pass through the corresponding point of the image.

principal focus of the lens L_1. In using a telescope the eye is usually adjusted for parallel rays; that is, it is focused for great distances. In that case, when focusing, one of the lenses is moved until I lies at the principal focus of L_2. The rays from one point of the image I, after emerging from the lens L_2, will form a beam of parallel rays.* These parallel rays are brought to a focus on the retina by the lens in the eye, and produce the same effect as if the point source from which the light comes were far away in the direction from which the parallel rays enter the eye. In the figure are shown the rays from two points of the source, which, for convenience, is supposed to be an arrow. The group of rays from each of these points forms a parallel beam after passing through the lens L_2. To the eye these points appear at a great distance away, and in the directions shown in the figure. In a similar way the light from other points on the object is focused on the retina of the eye. Hence the image appears far away and subtends the angle β of Fig. 426 or 427. *A telescope enlarges the angular size of the image seen by the eye.*

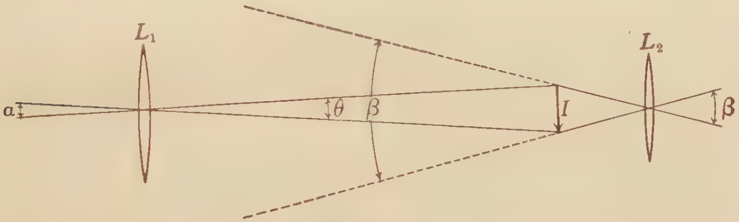


FIG. 427

The *magnifying power* of a telescope is defined in a different way from that of a simple magnifying glass; namely, in terms of the angular sizes of the image and of the object. In Fig. 427, α is the angular size of the object as seen by the unaided eye, and β the angular size of the image as seen by an eye placed close to the lens L_2.

---

*If the rays are parallel, they will be parallel to a line drawn from the source through the optical center of the lens. A line drawn in the figure from the arrowhead on I through the optical center of L_2 gives the direction of the parallel beam of the light which comes from the arrowhead.

The magnifying power of the telescope is

$$m = \frac{\beta}{\alpha} = \frac{\beta}{\theta},$$

since $\theta = \alpha$ by construction.

But in actual practice these angles are very small. Approximately, therefore (see equation (32) of section 77),

$$\theta = \frac{\text{size of image } I}{\text{distance from } L_1} = \frac{I}{F},$$

where F is the focal length of the lens L_1; and

$$\beta = \frac{\text{size of image } I}{\text{distance from } L_2} = \frac{I}{f},$$

where f is the focal length of the lens L_2.

Hence $$m = \frac{\beta}{\theta} = \frac{F}{f}. \qquad (27)$$

Hence the magnifying power of a telescope is the ratio of the focal lengths of the objective and the eyepiece lens. Therefore, to secure a large magnification the focal length of the objective must be large, and that of the eyepiece lens small. When the object is relatively near, equation (27) must be regarded as an approximation.

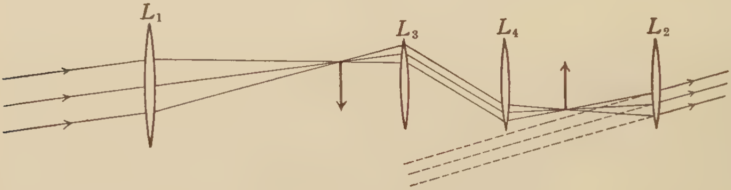


FIG. 428. The terrestrial telescope

624. The terrestrial telescope. The telescope described in the last section produces an inverted image. For many purposes this is objectionable. The terrestrial telescope, which is the kind used on many surveying instruments, is designed to give an erect image. In Fig. 428 are shown the extra lenses, L_3 and L_4, which are inserted to invert the image. How they do this may be learned from a study of the figure. Since the image produced by L_1 is formed at the principal focus of L_3, the rays emerging from L_3 are parallel. The second image is formed at the principal focus of L_4. When the eye is focused for parallel rays, this second image is at the principal focus of the eyepiece lens, L_2.

625. The opera glass. In the opera glass (or Galileo's telescope, as it is sometimes called) a *diverging* lens (Fig. 429) is used as the eyepiece lens. The converging objective lens L_1 would, if the lens L_2 were absent, form a real, inverted image at I_1. The diverging lens L_2, however, bends the rays as shown in the figure; hence, if the eye is placed close to the lens L_2, one sees an erect, virtual image, I_2. In the figure the image I_2 is shown, for convenience, much nearer than it usually is. The virtual image is a great distance away if the eye is focused for parallel rays, as is the case with other forms of telescopes. In focusing for this condition, the lens L_2 is moved until the image I_1 is at the principal focus of the lens L_2.

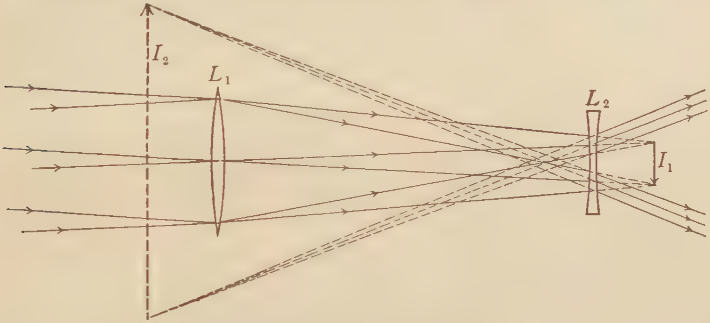


FIG. 429. The opera glass

The opera glass has a relatively small field of view and is suitable for only small magnifications. Even large-sized field glasses of this type magnify only from three to six times. Their chief advantage is that they are shorter than the telescopes described above.

626. The prism binocular. The best field glasses for giving good magnifying power and relatively large fields of view are the *prism* *binoculars.* In Fig. 430 is shown the ray path. This type is essentially an astronomical telescope with the inversion produced by reflection inside totally reflecting prisms. The rays, after passing through the objective lens L_1, are totally reflected back by the prism A to the prism B. This prism reflects the rays through the eyepiece lens L_2. The length of the optical path is nearly three times the length of the instrument, and permits the use of an objective with longer focal length, thus giving greater magnification.

The images are formed in the same relative positions as in the case of the astronomical telescope. If the eye is focused for parallel rays, the image produced by the objective L_1 is at the principal focus of the eyepiece lens. The two prisms are placed at right angles to each other. The reflection in one interchanges the two sides of the image, and the reflection in the second interchanges the top and bottom. Hence the image is seen in its proper position.

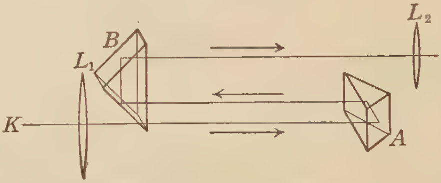


FIG. 430. The prism binocular

The term *binocular* arises from the fact that the telescopes are mounted in pairs, one for each eye.

627. The reflecting telescope. In the reflecting telescope a concave mirror is used to form a real image of the object. This image is viewed with an eyepiece lens as in the case of the astronomical telescope. Inasmuch as the light is reflected back, special arrangements must be made so that the observer does not obstruct the incoming light. In one method the mirror is tipped a little, and the reflected beam then forms the image off at one side; in another a small plane mirror or a reflecting prism is placed directly in front of the mirror and reflects the converging rays off to one side.

Very large reflecting telescopes have been constructed for astronomical purposes. A number have been built which are much larger than any refracting telescope. There is one at the Mt. Wilson Solar Observatory (near Pasadena, California) having a concave mirror 100 inches across. The largest lens in use is the objective in the large telescope of the Yerkes Observatory at Williams Bay, Wisconsin. This is 40 inches in diameter.

628. The range-finder. In Fig. 431 is shown the plan of one of the best types of range-finder. At a fixed distance apart are mounted two totally reflecting prisms of the type called *penta* prisms. In this type the two reflecting surfaces are at an angle of 45 degrees with each other, and the reflected ray always makes an angle of 90 degrees with the incident one, no matter what the direction of the incident ray may be, provided the ray enters and leaves the two faces which are at right angles to each other. This type of prism is sometimes called a *constant-deviation-reflecting* prism. Near each of the penta prisms B and D are the lenses L_1. Each of these lenses, together with the reflecting prism M or N and the eyepiece lens L_2, forms a telescope. The reflecting prisms M and N are placed one above the other; hence to one looking through the eyepiece lens the upper half of the field of view is seen by means of the light coming through the prism D, and the lower half by means of the light coming through the prism B. If the object viewed is at such a great distance that the rays A and C are parallel, and the reflecting faces of M and N are at right angles, the two halves of the field of view will match, and one continuous image will be seen. But since the two prisms M and N are not set exactly at right angles, a thin wedge of glass must be placed at W_1 to bend the rays so that the two halves of the field will match. If, now, a nearer object is looked at, the light rays from it will not be parallel, but will have the directions of the lines A and E. The ray leaving D will be bent downward, as shown by the broken line in the figure. In order to make the two halves match, the thin wedge must be moved to the position W_2. In a similar manner it can be shown that for different distances of the object the wedge must be moved to different positions. The instrument is so constructed that the motion of the

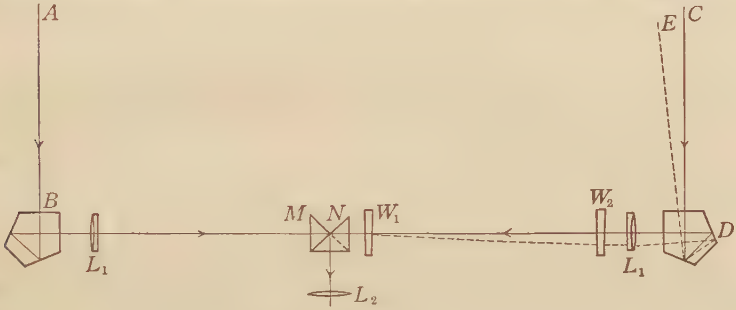


FIG. 431. The range-finder

wedge moves an index along a calibrated scale, from which distances of the object can be read off directly. For example, when the wedge is at W_1, the scale should read infinity; when it is moved a little, 10,000 yards; when it is in a position much nearer W_2, 1000 yards.

PROBLEMS

1. A photographic camera has a focal length of 6 in. How much farther from the lens is the image of an object 10 ft. away than that of an object 100 ft. away?

2. The picture on a lantern slide is $2\frac{1}{2} \times 3\frac{1}{2}$ in. What should be the focal length of a projection lens in order to make the picture on the screen $7\frac{1}{2} \times 10\frac{1}{2}$ ft.? (The screen is 30 ft. from the lens.)

3. The projection lens of a lantern has a focal length of 8 in. When the screen is 20 ft. from the lens, what will be the magnification of a slide?

4. If the distance of most distinct vision is 25 cm., what is the magnifying power of a converging lens with a focal length of 3 cm., when the lens is used as a simple microscope?

5. In a compound microscope the image formed by the objective lens, focal length 5 mm., is 20 cm. from the lens. The eyepiece lens produces a magnification of 10. Find the total magnification.

6. A compound microscope has an objective with a focal length of 4 mm., and an eyepiece lens with a focal length of 4 cm. The real image produced by the objective is 20 cm. from the objective. The final image is 25 cm. from the eyepiece lens. Compute the magnification.

7. A microscope having an objective lens with a focal length of half an inch is used to form an image of a slide upon a screen which is 20 ft. from the lens. What is the magnification?

8. The objective lens of an astronomical telescope has a focal length of 6 ft., and the eyepiece lens a focal length of 2 in. Compute the magnification that the telescope will produce when used for viewing distant objects.

9. The objective of an astronomical telescope has a focal length of 120 in., and the eyepiece lens a focal length of 2 in. While viewing a far-distant object the distance between the lenses was $121\frac{15}{16}$ in. (*a*) Where was the image that the eye saw? (*b*) Was the image real or virtual, erect or inverted?

10. The objectives of a pair of binoculars have focal lengths of 10 in. How much farther from the objective lenses are the images of an object 20 ft. away than those of an object very far away?

CHAPTER XLV

SPECTRA AND COLOR PHENOMENA

Pure spectra, 629. Types of spectra, 630. Emission spectra, 631. Absorption spectra, 632. Summary of types of spectra, 633. Luminescence, 634. Some applications of spectrum analysis, 635. The Doppler effect, 636. The complete spectrum, 637. Methods of detecting ultra-violet and infrared radiations, 638. The rainbow, 639. Coronas; halos, 640. The color of a body, 641. The mixing of colors, 642. Complementary colors, 643. The mixing of pigments, 644. Fluorescence; phosphorescence, 645.

629. Pure spectra. Newton showed by a method explained in section 598 that white light is a mixture of different kinds of light —kinds which produce different colors. It is frequently a matter of importance, not only in work in physical science but in many other cases, to be able to break up a beam of light into its component parts. The simplest method is to produce a spectrum by refraction through a prism. In doing this, one needs to employ certain precautions; otherwise he will not get a *pure* spectrum but will get, at each point in the spectrum, mixtures of different colors. In a pure spectrum there must be no appreciable overlapping of colors; that is, the light which reaches a given narrow region of the spectrum must be light of only one kind.

In the production of a pure spectrum a *narrow* source of light is absolutely necessary (the reason for this will appear later). In Fig. 432, S is a narrow slit, the length of the slit being perpendicular to the plane of the figure. As far as the focusing is concerned, the slit is the source of light. Let us suppose that the source emits only one kind of light, say blue light. When the lens L is placed so that the image of the slit will be at S', the prism P placed in the path of the light will cause the image to be formed at S_1. In this case there will be no breaking up of the beam of light by the prism, for light of only one frequency is present. Hence a sharp, narrow image of the slit S can be produced at S_1. But since a prism bends light different amounts for different frequencies, if another kind of light were present, say a green light, another image would appear at S_2 (green light has a smaller frequency and longer wave-length and is bent less than blue light). Under these circumstances two images of the slit—one blue, the other green—are seen. If the slit is narrow, these images are narrow lines; if the slit is wide, the images are also wide. If red light is present, it forms another image at S_3, and so on, *each additional kind of light forming its own image of the slit.* The images of those colors which have wave-lengths longer than the blue and shorter than the red are formed between S_1 and S_3. When the source is a white-hot body, like the filament of an incandescent lamp, the light is not made up of a few different kinds of light, but all kinds are present. Such a source emits all the waves which will affect the eye,—an infinite number of different frequencies. Hence, when such a source is used, the entire space between the image of the slit formed by the shortest wave that we can see and the one formed by the longest visible wave will be filled by an infinite number of slit-images. *Unless each of these images is very narrow*, there will be a considerable amount of overlapping. Hence, to get a pure spectrum a very narrow slit must be used, and the focusing must be very carefully done.

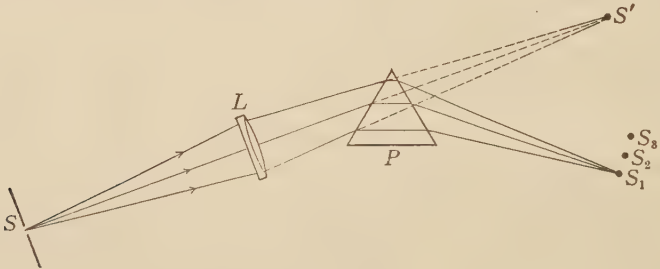


Fig. 432

In the method shown in Fig. 432 a serious difficulty arises if the spectrum is formed near the lens. In that case the beam of light leaving the lens L is converging, and strikes the prism P at different angles of incidence. Since the amount the light is bent depends on the angle of incidence, the light will not all be bent alike, and sharp images cannot be formed. It is important that all parts of the beam of light have the same angle of incidence on the prism. Only when the spectrum is focused on a distant screen, as is often done in a lecture experiment, is the method of Fig. 432 sufficiently accurate.

In Fig. 433 is shown diagrammatically an arrangement in which the entire beam of light makes the same angle of incidence on the prism. The lens L_1 is placed so that the slit S is at the principal focus; hence the rays emerging from L_1 form a parallel beam. The light emerging from the prism is still made up of parallel beams, but the parallel beam of red light is not parallel to the beam of violet light. In the figure the solid lines show the path of the red light, and the broken lines those of the violet light. The lens L_2 will bring these parallel beams to a focus in the focal plane of the lens. This is shown in the figure only for red and violet light.

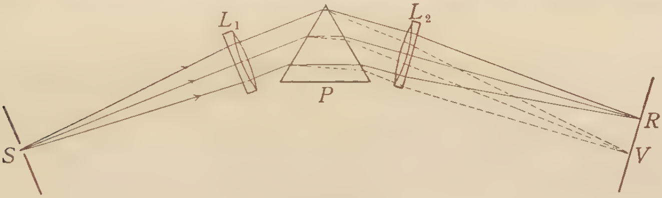


Fig. 433

Nothing has been said as yet about the position of the prism. For the best definition it should be turned until the incident rays and the emergent rays make nearly equal angles with the faces. This is called the position of *minimum deviation*, for the light is bent less than if the prism stands in any other position. This setting of the prism is done by trial, usually by turning the prism until the deviation is a minimum.

It appears from the foregoing that in order to get a pure spectrum: (1) an extremely narrow source must be used, (2) the light falling on the prism must be a beam of parallel rays, (3) the focusing must be carefully done, and (4) the prism must be set for minimum deviation.

630. Types of spectra. Spectra may be divided into two classes:

1. *Emission spectra.*
2. *Absorption spectra.*

The first of these may be divided into two subdivisions:

a. Bright-line spectra.
b. Continuous spectra.

The second class—the absorption spectra—may be divided into two subdivisions:

a. Dark-line spectra.
b. Dark-band spectra.

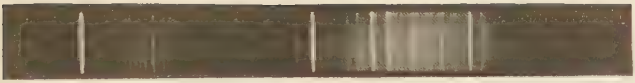
631. Emission spectra. Incandescent or luminescent vapors and gases, except when under great pressure, give *bright-line* spectra.

Incandescent solids and liquids, and gases or vapors under great pressure, give *continuous* spectra.

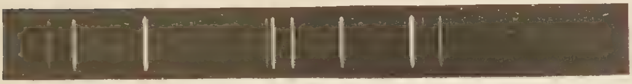
If an electric discharge is passed through a gas at low pressure contained in a vacuum tube, the gas becomes luminous and gives a light which when analyzed shows a spectrum made up of a number of narrow bright lines. The spectrum of any one luminescent gas is different from that of any other gas, not only in the position of its lines but in their number; hence it is often a simple matter to identify the kind of gas by its spectrum.

A photographic reproduction of the bright lines in the spectra of several gases is given on the plate facing this page. That from mercury was obtained from a mercury arc lamp. In the other cases the gases were at low pressures in glass tubes and were made luminous by an electrical discharge. Such photographs do not bring out the great beauty of these spectra, since the colors are not reproduced.

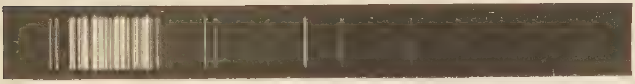
If a little common salt is added to a Bunsen flame, which is usually nearly colorless, the flame is filled with the vapor of the salt; a bright orange light is given out. The spectrum of this light consists of two bright lines so close together that they blend unless the instrument used gives considerable dispersion (color separation). Further experiments show that this light is due to the sodium in the salt. For most purposes the light from a "sodium flame" may be regarded as monochromatic; that is, having only one frequency.



Hydrogen



Helium



Neon

Argon

Mercury

BRIGHT-LINE SPECTRA

In the light from white-hot solids, such as the hot carbon of an electric arc light, the filament of an incandescent lamp, or the incandescent carbon particles in a candle flame or a gas flame, all colors are present; hence the spectrum formed from the light of such sources is *continuous*.

A little thought will enable the student to see why we might expect a luminous gas to emit only a relatively small number of different frequencies, and a solid to emit all frequencies. Undoubtedly the light is produced by the vibrations of the electrons in the atoms. In a gas the atoms are relatively far apart, and the electrons have a chance to vibrate in their natural modes without interference from neighboring atoms; but when they are packed, as in a solid, there may be considerable interference. The strings of a musical instrument give out definite tones only when they are free to vibrate in their natural modes: but if various instruments are crowded together so that the strings touch here and there, tones other than the natural ones are emitted. Apparently in solids all modes of vibration of the electrons, within wide limits, become possible; hence all frequencies are emitted.

632. Absorption spectra. When white light passes through a sheet of red glass, violet, blue, and green may be completely absorbed. Blue glass may absorb all the light which forms the red end of the spectrum. A dilute solution of potassium permanganate absorbs the middle region of the spectrum. Many solids and liquids thus absorb groups of waves and transmit other kinds. Hence, in the spectrum of white light which has passed through such absorbing media, wide *bands* are absent. Light which has passed through such material forms what is called a *dark-band spectrum*.

But when white light passes through a gas or vapor, the absorption that takes place is different. Instead of a dark band in the spectrum, narrow dark *lines* appear, and only very narrow regions are absorbed. A well-known and very famous experiment is the absorption produced by sodium vapor. A spectrum of white light from the carbon of an arc light is projected on a screen. In the path of the light between the source and the slit is placed a Bunsen flame. When a little metallic sodium is held in this flame in a small iron spoon, and is vaporized, the flame becomes filled with an orange-colored vapor. The absorption of light by the sodium vapor produces a *narrow black line* in the orange part of the spectrum of the arc. If, while the flame is luminous, the arc light is extinguished, a faint orange line—the spectrum of sodium—will appear on the screen. It will be noticed that the *bright line formed by the vapor is in exactly the same position in the spectrum as the absorption line of that vapor.*[*]

This is only one example of a general rule: *Heated vapors and gases absorb light of the same frequency as they emit.*[†]

Since incandescent vapors and gases produce bright-line spectra when emitting light, they must, when absorbing, produce dark-line spectra.

The spectrum of the sun is a dark-line spectrum showing hundreds of fine lines. Since it is an absorption spectrum, we are certain that sunlight has passed through some absorbing medium. As the absorption is in the form of dark *lines*, the absorbing medium must be gaseous. The absorbing medium is the atmosphere of the sun and that of the earth. In fact, this is one of the ways by which it is found that the sun is surrounded by vapors.

If the dark-line spectrum of the sun and the bright-line spectrum of iron vapor are placed one above the other so that they match (that is, that light of any frequency in one lies immediately above the same frequency in the lower), it will be seen that for each bright line in the iron spectrum there is a dark line which has the same position in the solar spectrum. In this way several hundred of the dark lines of the spectrum of the sun have been identified as due to the absorption of iron vapor. Hence we are certain that there is iron vapor in the atmosphere of the sun. The temperature of the atmosphere of the sun is so high that many different metals exist there as vapors. Many of these vapors have been detected by the position of their absorption lines.

The dark lines in the solar spectrum are called *Fraunhofer lines*, because Fraunhofer was the first to make careful measurements of them, although he was not the first to observe them. The most conspicuous Fraunhofer lines are named by the letters of the alphabet. Modern instruments show that some of these named lines are really groups of many lines. Thus the *A* and *B* lines are groups of fine lines in the red, produced by the absorption of the earth's atmosphere; the *C* line, also in the red, is due to hydrogen in the sun's atmosphere; the *D* line is a double line in the orange, caused by sodium vapor in the sun's atmosphere.

---

[*] If an instrument of large dispersion is used, the sodium "line" will consist of two lines very close together.

[†] See section 634, on luminescence.

633. Summary of types of spectra. *Bright-line* spectra are produced by incandescent or luminescent vapors and gases.

Continuous spectra are produced by incandescent solids or liquids, or by gases under great pressure.

Dark-line spectra are produced when white light from a solid or a liquid, or a gas under enormous pressure, passes through absorbing vapors.

Dark-band spectra are produced when white light passes through an absorbing solid or liquid.

634. Luminescence. Gases and vapors do not always absorb light of the same frequency that they emit. A gas or vapor must absorb the same frequencies it emits, provided that the gas or vapor is at the same temperature when absorbing as it is when radiating, and provided also that the radiation is caused by heat. But there are cases of radiation other than those produced by heat; for substances may radiate light without being hot, as in the case of fluorescence and phosphorescence (sect. 645). The radiation from the electric discharge through a gas in a vacuum tube is not a high-temperature radiation. Some of the nebulæ and comets are probably not at a high temperature, although they emit light.

Nearly all vapors have bright emission lines for which there are no corresponding absorption lines, but there is always a bright emission line corresponding to each absorption line of a substance.

The term *luminescence* is applied to those forms of radiation which apparently are not due to high temperature.

635. Some applications of spectrum analysis. The spectra of all the different types of bodies seen in the sky have been carefully studied, with the result that a vast amount of knowledge of the heavenly bodies has been acquired. Since the planets and our moon give identically the same spectrum as the sun, their light must be reflected sunlight. The light from the edge of the sun, best seen at the time of a total eclipse, gives a bright-line spectrum. The bright-yellow line of helium gas was found in this spectrum long before the element was discovered on the earth. The spectrum of the light from the edge of the sun thus shows beyond all doubt that the sun is surrounded by an atmosphere of incandescent vapors.

Those bodies which are found to give bright-line spectra are at once known to be gaseous bodies. Thus, many nebulæ and comets are known to be gaseous. Most of the stars give dark-line spectra similar to that given by our sun, showing a similar structure,—a vast heated mass surrounded by hot vapors. Some of the stars show some bright lines superposed on a dark-line spectrum, indicating that some of their light comes from gases.

Much of the evidence used in the theories of the evolution of the nebulæ and stars comes from a study of the spectrum. In the next section is given still another method of using the spectrum in astronomical work.

The study of the spectra of various substances has been of great service in many chemical investigations. It was of great assistance in the isolation and discovery of the rare elements in the atmosphere: helium, neon, argon, krypton, and xenon. Cæsium, rubidium, and thallium were discovered by finding new bright lines in the spectra of certain substances.

636. The Doppler effect. It was pointed out in section 331 that when a source of sound is moving toward an observer, or the observer is moving toward the source, the apparent frequency of the sound is increased and the wave-length decreased; when the source and the observer are moving apart, the frequency is decreased and the wave-length increased. The same reasoning applies to light-waves: we should expect that light from a star which is moving away from us would show a lengthening of its waves and a decrease in frequency, and that one moving toward us would show a shortening of the waves and an increase in frequency. Since most of the spectra of the stars show dark lines the wave-lengths of which are known, it is relatively easy to find out whether these lines are shifted toward the red or toward the violet end of the spectrum. If they are shifted toward the red end, the waves are lengthened, and the star is moving away from us; if they are shifted toward the violet end, the wave-lengths are shortened, and the star is moving toward us. By measuring the magnitude of the shift of the position of the lines of a star spectrum, it is possible to compute both the change in wave-length and the velocity of the star toward us or away from us. This gives only a component of the star's velocity — the velocity in the "line of sight."

There is a very interesting application of this principle in the case of some double stars. It seems that there are many cases in which a star is in reality a pair of stars, relatively close together, which revolve around each other. If we could see separately the stars in a pair which rotate in an orbit in a plane nearly in the line of sight, at one time we should see one star moving toward us, and the other away from us. A few hours or a few days later one of the stars would have moved to the front, and the other around behind the first. At that time neither star would be moving toward or away from us. A little later we should see both — the first moving away, and the second toward us. But we cannot see this with the eye, not even with the most powerful telescope. We know this is what is happening, for two reasons: in the first place, the intensity of the light changes periodically, provided the stars are moving so that one goes behind the other; the second and more important reason is that we observe periodic changes in the star's spectrum. At one time each spectrum line is doubled; at another time the lines have moved together and coincided in single lines. The doubling of the lines is due to the fact that when one star is moving toward us and the other away, the lines produced by one are shifted toward the violet end of the spectrum, and the lines of the other are shifted toward the red end. Hence each line is broken up into two. At another time neither star is moving toward or away from us, and there is no change in the wave-lengths: the lines are single and in their normal positions.

637. The complete spectrum. The eye is sensitive to only a small part of the radiation from hot substances. The shortest wave that affects the eye produces the sensation of violet, while the longest wave produces the sensation of red. Waves which are shorter than the violet are called *ultra-violet* waves, and those longer than the red, *infra-red*.

The following table gives an approximate idea of the range of wave-lengths that have actually been detected and measured in the radiation from sources which give out light:

Shortest ultra-violet, about	.000002 cm.
Limit of visible violet	.000038 cm.
Limit of visible red	.000076 cm.
Longest infra-red, about	.03 cm.

Most of the radiant energy of all heated solids and liquids is in the infra-red (see also section 253).

638. Methods of detecting ultra-violet and infra-red radiations. The photographic plate is sensitive throughout the entire range of the ultra-violet region of the spectrum. By special treatment of the emulsion, these plates have been made sensitive to waves which are considerably longer than those of red light; but in general the methods which have been used for detecting infra-red radiations are based on the conversion of the energy in these radiations into heat.

639. The rainbow. The most commonly seen spectrum is the rainbow. It is not a pure spectrum, such as is explained in section 629; for there is much overlapping of colors, due chiefly to the fact that the source, the sun, is not narrow. The rainbow is seen only when the sun is low in the sky and the observer is looking away from the sun. The *primary* rainbow is usually accompanied by a faint *secondary* lying outside of it. The angular radius of the primary is about 41 degrees, and that of the secondary is about 52 degrees. Close to the inner edge of the primary and close to the outer edge of the secondary are faint supernumerary, or complementary, bows. In the primary the violet is on the inside, and the red on the outside; in the secondary the order of colors is reversed. The reason for this color distribution will appear later.

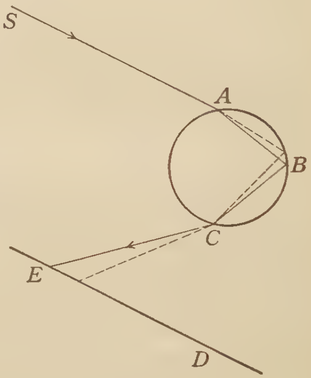


FIG. 434

In Fig. 434 the lines *SABCE* show the path of a ray of light which is reflected once inside a drop of water, and refracted as it enters and as it leaves. In Fig. 435 the lines *SABB'CE* show the path for the case of two internal reflections.*

In both these figures the solid line shows the path of violet light, and the broken line shows the path of the red light, which is bent less than the violet. The angular separation of the violet and the red has been exaggerated in both figures. In both cases the lines *ED* are parallel to the incident beam. In the first case the angle *CED* for the violet light should be 40° 36′, and for the red light 42° 22′; in the second case the angle *CED* should be 53° 36′ for the violet light, and 50° 24′ for the red.

---

*The rays drawn in the figures are those for minimum deviation. For other angles of incidence the light reaching the eye is absent or faint.

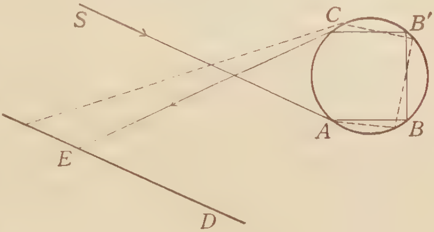


FIG. 435

Imagine a beam of sunlight shining upon a large number of raindrops (see the simplified diagram in Fig. 436). Through the eye at *E* is a line *ED* drawn parallel to the incident sunlight. Those drops which are in the direction *EA* will refract violet light to the eye if the angle *AED* is 40° 36′; those in the direction *EB*, if at an angle of 42° 22′, will refract red light; while the drops in directions between these will refract other colors to *E*. Hence, if the eye were at *E* and looking toward *B* and *A*, a spectrum band would be seen. Similarly, at a somewhat greater angle, one sees the spectrum of the secondary bow.

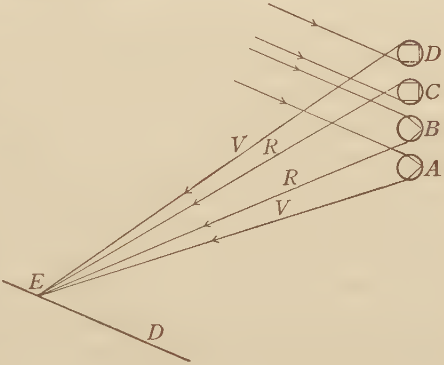


FIG. 436

Since the sun is far away, the line *ED* may be said to be a line drawn from the sun through the observer's eye. Drops in any direction that makes an angle of about 40 degrees with the line *ED* should refract violet light to the point *E*; from any direction making an angle of about 42 degrees, red light should be refracted to the eye; and so on. Hence if the figure is rotated about the line *ED* as an axis, *AB* traces out an arc, giving the direction of the primary bow, as seen from *E*, and *CD* traces out an arc giving the direction of the secondary bow. When the sun is near the horizon, one can often see bows making almost complete semicircles; but when the sun is higher in the sky, a smaller portion of them can be seen. (What is the greatest angular height of the sun at which a part of a primary rainbow can be seen?)

640. Coronas; halos. *Coronas* consist of one or more bright-colored rings of relatively small diameter surrounding the sun or moon, seen when looking through a fog or cloud of fine water drops. One who looks through mist-covered glasses at a bright but small source of light will usually see these colored rings. They are due to diffraction (the bending of light-waves around small obstacles) and to interference (Chapter XLVI).

Halos are the larger circles often seen around the sun and moon. The best-known occur in two sizes: one having an angular radius of 22 degrees, and the other a radius of 46 degrees. They are caused by refraction of light through ice crystals floating in the air. The cirrus clouds and others formed at temperatures below 0° C. consist of small snowflakes or ice spicules. These tiny crystals are always formed with faces making angles of 60 and 90 degrees. When light passes through an edge of an ice crystal, it is going through a prism of either 60 or 90 degrees. The 22-degree-radius halos are produced by the light which is refracted by the 60-degree angles, and the 46-degree-radius halos by the light which is refracted by the 90-degree angles. The latter kind is not so frequently seen as the former, but occasionally it may be very bright.

Parhelia, or *sun dogs*, are produced by the light refracted through the ice prisms of 60-degree angles. They are seen at an angular distance of 22 degrees on each side of the sun when the sun is near the horizon and when the air contains large numbers of ice crystals.

641. The color of a body. If one projects a spectrum of white light on a screen, and then places in various parts of it a red rose or a piece of red cloth, he will find that the red object is black or brown in all parts of the spectrum except in the red. Such an object can reflect only red light; it absorbs all other colors.

Similarly, a piece of green yarn reflects well in the green, but in the blue or red part of the spectrum it fails to reflect. These experiments explain the common cause of the color of a body. When white light falls on an object which reflects the red component and absorbs the other colors, the object must appear red; and similarly with objects of other colors. The color of a body is caused by *selective reflection.*

The observed color of a body depends also on the kind of light with which it is illuminated. In the light from a sodium flame or from the mercury arc there is practically no red present, and a red object illuminated only by this kind of light cannot appear red. Nearly everyone knows that there is often great difficulty in matching two samples of cloth by artificial light. The two may appear to be the same in a yellowish light, but when seen in a good white light, such as sunlight, they appear very different.

The color of red glass is due to absorption. It transmits only red light, absorbing the rest. The colors of other colored glasses and liquids are similarly explained.

642. The mixing of colors. The colors seen in a pure spectrum are pure colors, for each of them is due to only one kind of light. There are many other colors which are not seen in the spectrum; purple, for example. The light transmitted by a weak solution of potassium permanganate is purple. If this solution is placed in the path of white light, the spectrum will show that the solution absorbs the central portion of the spectrum and transmits the violet and the red; hence purple may be regarded as a mixture of these two kinds of light. All known colors can be resolved in terms of the colors which are seen in the spectrum; in other words, any color can be produced by the proper mixture of spectrum colors. Only the colors seen in the spectrum are pure; all others are mixtures.

643. Complementary colors. The spectrum shows that white light from a hot source is a mixture of all the spectrum colors. When these colors are added together again, they form white light. But not all the spectrum colors are needed to produce a light which gives the sensation of white; for example, this may be produced by adding blue and yellow lights. Any two colors which when added together produce the sensation of white are called *complementary* colors.

644. The mixing of pigments. The mixing of colored lights must not be confused with the mixing of colored paints or pigments. To mix blue and yellow lights they are thrown on a white screen together. The screen *reflects both blue and yellow* to the eye. But when blue and yellow pigments are mixed, the process is very different. A pigment is blue because it absorbs many of the other colors; the common blue pigments absorb entirely red and yellow and reflect blue and some green. A yellow pigment absorbs the colors at the violet end of the spectrum, and reflects yellow, some red, and also some green. When the pigments are mixed, the blue pigment absorbs red and yellow light falling on the mixture, and the yellow pigment absorbs violet and blue. Green is the only color reflected by both. Hence the mixture is green.

There is a simple experiment which involves the same principle. If a glass lantern cell is filled with a dilute solution of potassium bichromate, the light transmitted will be yellow. Another cell, preferably a thicker one, is filled with a dense solution of copper sulfate. Blue light is transmitted by this cell. Hence we have one cell transmitting yellow light, and the other blue light. Now if these two cells are held so that they overlap, it will be seen that the color transmitted by the two is *green*. The explanation is as follows: The yellow liquid transmits some green but no blue or violet, as can be shown by the spectrum of the transmitted light. The blue solution transmits some green but no yellow or red. Green is the only color transmitted by both.

645. Fluorescence; phosphorescence. If a piece of canary glass is placed in the violet end of the spectrum, it will give out a greenish color, although no green light has fallen on it. In some way it is able to absorb the short waves (the violet or ultra-violet) and to emit this energy in a lengthened wave—a green. Such a phenomenon as this is called *fluorescence*.* Short X-rays falling on certain chemicals are absorbed, and the energy is emitted in a radiation of longer wave-lengths, one that we can see. The reason why fluorescent screens are so useful in X-ray work is obvious: they convert an invisible radiation into a visible one. Many substances show fluorescence; nearly all petroleum products, such as kerosene, vaseline, and lubricating oils, fluoresce with a bluish glow.

---

*A law due to Stokes states that the emitted fluorescent radiation always has a longer wave-length than the absorbed radiation.

Fluorescence is something different from reflection. When light is reflected, there is no change in the frequency or the color. When a body fluoresces, it absorbs light of one frequency and emits a different kind, usually one having a lower frequency or a longer wave-length.

In some cases fluorescence persists after the exciting source of light has been removed. In this case it is called *phosphorescence*. Sometimes it persists for only a fraction of a second, while with other substances the phosphorescent light may be given out for hours afterwards. Phosphorescent paints can be purchased which will absorb light during the daytime and reëmit it at night. This must not be confused with the luminous products containing radium which are now coming into such common use, nor with the slow oxidation of phosphorus, as in the glowing trail left by a damp match.

Fluorescence and phosphorescence are special cases of the more general subject of luminescence.

CHAPTER XLVI

INTERFERENCE AND DIFFRACTION

Introduction, 646. The interference of waves, 647. The interference of light-waves, 648. The Fresnel mirrors, 649. The Fresnel biprism, 650. The colors of thin films, 651. Newton's rings, 652. Interference and the wave theory, 653. Diffraction, 654. Some experiments showing diffraction of light, 655. The diffraction grating, 656. The theory of the diffraction grating, 657. An alternative explanation of the grating, 658. X-ray spectra, 659.

646. Introduction. The subjects of interference and diffraction of light may be looked at from two different standpoints, both of them important.

1. A study of these subjects furnishes definite evidence that light is propagated by wave-motion. Recently the same method of reasoning has been used to prove that X-rays also are propagated by a wave-motion. The study of interference and diffraction gives us methods by which it can be determined whether any radiation is a wave-motion or not.

2. There are numerous applications of the principles which are stated in this chapter. The measurements of the wave-lengths of light and X-rays, and the measurements of other extremely short distances, are based on interference. The most refined measurements which have ever been made in many different fields are based on the interference of light.

647. The interference of waves. Attention has been called to the fact that when the paths of two or more wave-trains cross, the disturbance at any instant is the sum of the disturbances of the different waves. Sometimes (see section 329) these disturbances add in such a way that the resultant at certain points is always zero. For example, suppose that two pieces of wire are attached to one prong of an electrically driven tuning-fork, and that these wires dip down into water or mercury. When the fork is vibrating, two different sets of circular ripples travel out from the wires, one set having one wire as center, and the other set having the other wire as center. Fig. 437 represents diagrammatically a portion of the disturbed region. The full lines show the position of the crests of the waves at one instant, and the broken lines the position of the troughs at the same instant. The wave-lengths of the two sets of waves are equal, because the two sources have the same frequency of vibration. Along the line A the crests and the troughs from the two sources meet *in phase*. They do this because each point on this line is equally distant from the sources S_1 and S_2. Along this line the disturbances add so as to reënforce each other. But all points on the line B are half a wave-length nearer S_1 than S_2; hence the disturbances from S_1 reaching any point on this line are half a wave-length ahead of those from S_2, and the two groups meet with opposite phases: crests are meeting troughs. Along this line at all times there is destructive interference. If the amplitudes of the waves given off by S_1 and S_2 are equal, there will be no disturbance along this line. All points on the line B' are half a wave-length nearer S_2 than S_1, and there is interference along this line. However, all points on the lines C and C' are one wave-length nearer one source than the other. Hence the waves meet in phase, as shown in the figure, thus reënforcing each other.

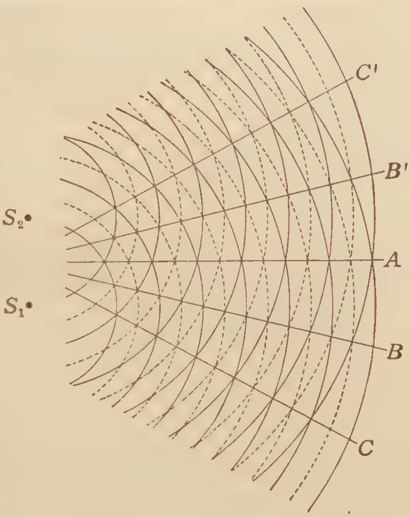


FIG. 437

In determining whether the waves reënforce or destructively interfere at any one point, the *path difference* of the two wave-trains must be considered. In the explanation of such cases as that given in Fig. 437, there is reënforcement when the path difference is zero, or one wave-length, or some whole multiple of a wave-length; but there is interference when the path difference is half a wave-length, or one and one-half wave-lengths, or two and one-half wave-lengths, or any odd number of half-wave-lengths. This statement may be regarded as a general rule.

The principles involved in this experiment are true for all types of waves, but the apparatus required and the details of the necessary adjustments are very different for different kinds of waves.

648. The interference of light-waves. There are two reasons why the experimental methods for showing interference of light are difficult:

1. Light-waves have extremely short wave-lengths; therefore a difference of path of half a wave-length is a very small distance.

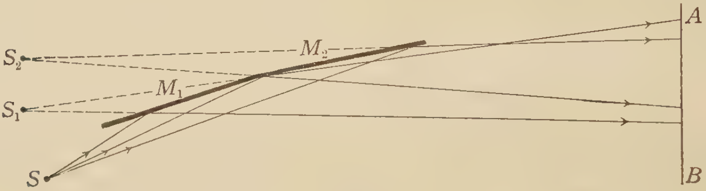


FIG. 438

2. In the case of the waves of Fig. 437 the two sources were made to vibrate together by attaching the two wires to the same prong of a tuning-fork. But it is impossible to get two different sources of light which will vibrate in phase. This difficulty in the case of light-waves restricts the experimenter to a single source.

In 1816 and 1826 Fresnel designed two different experiments which successfully meet these difficulties and which show clearly the interference of light-waves.

649. The Fresnel mirrors. Fresnel used two mirrors, M_1 and M_2 (Fig. 438), to reflect light from a narrow source S to a screen AB. The light that is reflected to the screen by M_1 appears to come from S_1 (the image of S in the mirror M_1), and the light reflected from the mirror M_2 appears to come from the image S_2. When some suitable screen (not shown in the figure) is so placed that no light reaches the screen AB directly from S, AB is illuminated only by light from two virtual sources, S_1 and S_2. The two sources S_1

and S_2 vibrate in phase, for they are both images of the same source; thus the second difficulty mentioned in the last section is overcome. By adjusting the mirrors M_1 and M_2 until they are nearly parallel, the images S_1 and S_2 are moved until they are very close together. When all the necessary adjustments are properly made, a series of narrow parallel bright and dark bands appear on the screen. These bands are perpendicular to the plane of the figure and parallel to the narrow source S. The central bright band is equally distant from the sources S_1 and S_2. The first dark band on either side is half a wave-length nearer to one source than to the other, the second dark band is 3 half-wave-lengths nearer, the third is 5 half-wave-lengths nearer, and so on. The path differences for the bright bands on each side is 1 wave-length, 2 wave-lengths, 3 wave-lengths, and so on.

Fresnel's apparatus can be used to measure wave-lengths of light. The details will be found in more advanced treatises, but the general principles can be understood from the following procedure: The angle between the two mirrors is measured, usually by some optical method based on reflecting light from the two mirrors. When this angle and the position of S with respect to the mirrors are known, the position of the two images, and their distance apart, can be computed. Then the following data are known or can be determined: (1) the positions of S_1 and S_2 with respect to the screen and (2) the distance between the bands on the screen. It is then easy to compute how much nearer one dark band is to S_1 than it is to S_2. If the band chosen is the first one, this difference in distance is half a wave-length; if it is the second, the difference is three half-wave-lengths. From this difference of path the wave-length is obtained. The method does not give accurate results, because one cannot measure the distance between the dark bands with a high degree of accuracy; but results correct within 5 or 10 per cent can be obtained.

650. The Fresnel biprism. Fresnel devised another and, in many respects, a more simple method of showing interference of light. A piece of glass was made with a double bevel on one side (P, Fig. 439). A plate of this kind is called a *biprism.* Light going through half of it is bent in one direction, and that going through the other half is bent in the opposite direction. The light from the slit S, after passing through the upper half of the biprism, appears to come from S_2; and that through the lower half, from S_1. Hence the light falling on the screen apparently comes from the two sources S_1 and S_2. If the biprism has been ground so that the images S_1 and S_2 are separated by a very small distance, the interference bands are clearly visible.

When a white-light source is used in either of the Fresnel methods, the bands will appear colored. In the method of the mirrors this is due to the fact that the wave-lengths of the different colors are not the same; hence a dark band for red is not in the same place as the dark band for a different wave-length. The same thing is true for the biprism method, but in this case there is the additional fact that the prism bends different colors different amounts. Hence the positions of the images S_1 and S_2 will be slightly different for different colors.

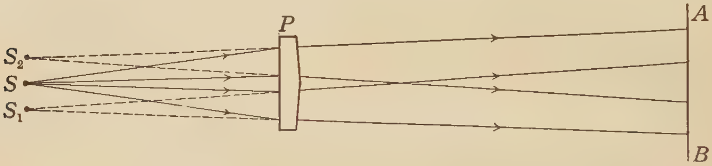


Fig. 439

651. The colors of thin films. All of us are familiar with the colors seen in thin films. Soap bubbles show beautiful tints. Glass can be blown out into thin-walled bubbles which show colors. A thin film of oil on water, or the thin layers of air sometimes seen in ice and other substances having a crystalline structure, show color effects. The colors seen in pearls are due to reflection from thin air films. In all these cases the colors are produced by interference of light.

In Fig. 440 the pair of lines AB represent on an exaggerated scale the two surfaces of a thin film which is slightly wedge-shaped. They may be the two sides of a soap film, or the two sides of a film of air between two glass plates, or any other sort of thin transparent film. S is the source of light; it may be some point on a large source of light. As the film is slightly wedge-shaped, the light which comes through a point P on the surface of the film travels from the source to the eye by two different paths. In one of these paths the light is reflected from the top of the film; in the other, from the lower side. If the eye is focused on the point P, the rays shown in the figure will reach the same point on the retina of the eye. As these rays have traveled different distances, they may not meet in phase on the retina. If the source emits white light, there may be interference for red light; that is, the beams traveling in the two paths through the point P meet with opposite phases and destroy each other. Since blue light has a different wave-length, two rays of blue light might not meet with opposite phases, but with the same phase, and thus reënforce each other. The destruction of red light and the reënforcement of blue would happen if the retardation* of one beam behind the other were equal to half a wave-length of red light or to some odd multiple of half a wave-length, and equal to a whole wave-length of blue light or to some multiple of a whole wave-length. The eye sees color in the film whenever waves interfere and cause one or more colors to be absent.

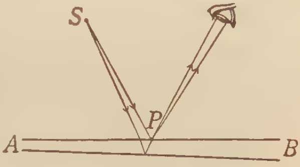


FIG. 440

When a thin film is viewed by reflected monochromatic light (by light from a sodium flame, for example), the film appears crossed by a large number of bright and dark bands. This can easily be seen if a piece of plate glass is laid on the top of another piece, and if a sodium flame is so placed that light reflected from the glass plates reaches the eye. In this case the film is the thin layer of air between the two plates, and the interference is between the two beams of light which are reflected from the two surfaces of glass bounding the air film. On account of the fact that the surfaces of the glass are not perfectly flat, and because small dirt particles accidentally get in, the air film will not be of uniform thickness. For one place on the film one of the reflected beams will be an *odd* number of half-wave-lengths of sodium light behind the other reflected beam, with the result that there is destructive interference. At another place the film may be thinner or thicker, and thus one beam may be an even number of half-wave-lengths behind the other. In that case there is reënforcement of light.

---

* It is explained in advanced treatises on light that at reflecting surfaces there are complicated changes of phase in light-waves. This effect produces, in the case of light reflected from the two surfaces of a film, the same difference of phase that would be produced by a path difference of half a wave-length. Hence the effective retardation of one ray behind the other is not produced solely by the difference in length of their paths.

If the two glass surfaces were perfectly flat, and if the film were made wedge-shaped by stray dirt particles or by a small piece of thin paper inserted at one edge, the light and dark bands would form smooth continuous lines; but if there were a slight irregularity in the surface of the glass, such as a slight depression in one surface, the bands would be distorted. Depressions of a depth equal to a fraction of a wave-length of light can be detected in this way. This is a very common and simple method for testing surfaces of glass for flatness. There are many optical instruments in which the surfaces of prisms or of other pieces of glass must be "optically" flat. Fig. 441, *a*, is a photographic reproduction, about natural size, of the interference bands formed by monochromatic light when an optically flat piece of glass was placed on the face of a high-grade glass prism. Fig. 441, *b*, shows the form the bands took when the same test plate was placed on a piece of plate glass. The appearance of the bands indicates that

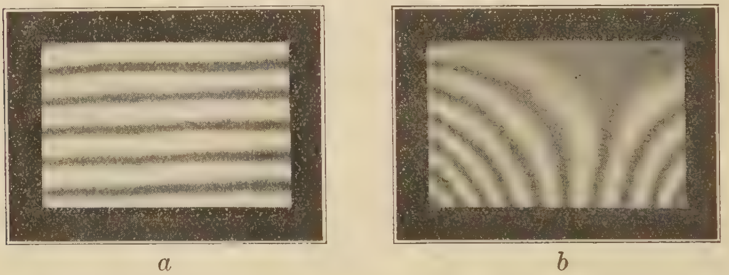

a *b*

Fig. 441. Interference bands in an air film between two glass surfaces. (*a*) The bands are nearly straight, showing that neither surface has a depression greater than about 1/500,000 inch. (*b*) One surface is that of ordinary plate glass

the surface of the plate glass was not optically flat. In the case illustrated in Fig. 441, *a*, the bands are nearly straight, showing that a band is at no place shifted more than about one fifth of the distance between adjacent bands. Since the distance between bands corresponds to a change in thickness of the air film of half a wave-length, this shows that neither surface contains depressions which have a depth greater than about one tenth of a wave-length (0.000006 cm.).

652. Newton's rings. Sir Isaac Newton found that when a convex-glass lens of large radius of curvature was pressed against a piece of plane glass, rings of colored light were seen around the point of contact when white light was reflected from the surfaces. This is another case of the interference of light from thin films. In this case the film of air between the two glass surfaces is not of uniform thickness. Regions of equal thickness of the film lie in concentric circles around the point of contact of the surface of the lens and the plane glass. Since the kind of color seen depends on the thickness of the film, the colors form concentric circles around the contact point.

653. Interference and the wave theory. Newton tried to explain the colors of thin films on the basis of the emission theory of light, but with poor success. After the more definite experiments of Fresnel, attempts to explain the observed phenomena by methods other than the wave theory were abandoned. Not only does the wave theory satisfactorily explain the facts, but there is a complete *quantitative* agreement. In recent years there have been developed many ingenious but intricate methods of producing interference by which either very accurate measurements of wave-lengths or accurate measurements of small distances can be made. The numerical results obtained by different methods check within the limits of experimental errors. No other satisfactory method of explaining interference phenomena has been developed, and there are reasons for believing that there never will be. It is now generally accepted that whenever interference can be demonstrated for any type of radiation, then that radiation is a wave-motion. Recently it was shown that interference of X-rays exists. Immediately all other theories of X-rays were abandoned in favor of a wave theory.

654. Diffraction. When waves meet an obstacle, they usually bend around it. Sound-waves bend around buildings, and water-waves bend around piers, as illustrated in Fig. 442. This bending around obstacles is called *diffraction.*

For a long time the apparent absence of diffraction of light was an argument against the wave theory. It was argued that since light travels in straight lines, light could not be a wave-motion; that if it were, light would bend around corners. In this connection attention should be called to the two following facts:

1. A more complete theory of wave-motion shows that long waves bend around corners to a much greater extent than short waves. This is confirmed by experiment. For example, if the student has ever stood on a cross street at some little distance from a street along which a band was marching, he probably was impressed with the change in the quality of the music as the band came into view. When the band was around the corner, the notes of the bass instruments—those that gave off the longer waves—were relatively much louder than those from the higher-pitched instruments; but as the band came into view, there was a rather sudden increase in the intensity of sound from the higher-pitched instruments—those that gave off shorter waves. This effect is caused by the fact that long waves bend around corners much more readily than short waves do. It is not difficult to show by experiment that relatively small obstacles cast "shadows" of the sound emitted from high-pitched sources.

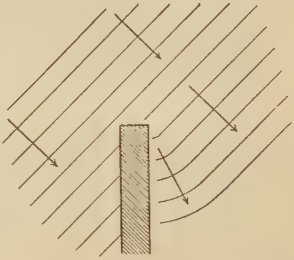


Fig. 442

2. More careful observations show that light does bend around corners. The smallness of the observed effect is readily explained by the shortness of the waves. Several experimental methods of showing these effects are described in the following sections.

655. Some experiments showing diffraction of light. The bending of light around the edges of objects is not commonly seen because ordinary sources of light are too wide. The shadows cast by sunlight or by light from most of our common sources have blurred edges on account of the width of the source. But when a narrow source of light is used, sharply outlined shadows should be obtained if light traveled strictly in straight lines. When an experiment with a narrow source is tried, it is found that light does bend, and hence does not form the expected sharp, or "geometrical," shadow. Some electric street lamps have short, narrow filaments. The student should try to observe the shadows of leaves and twigs cast on a sheet of paper by such a source. The source should be at some distance—40 feet or more—beyond the leaves or twigs, and the shadow observed should be 20 feet or more from the leaves or twigs. A distinct penetration of light into the geometrical shadow can be seen.

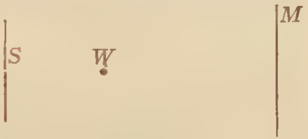


Fig. 443

When a bright source of light* is placed behind the narrow slit *S* (see Fig. 443), in such a position that a shadow of a fine wire *W* is formed on the screen *M*, some interesting things are observed. (The wire must be parallel to the slit *S*; in the figure both are supposed to be perpendicular to the plane of the paper. The wire should be 5 or 6 feet from the slit, and the screen some 10 feet farther.) An inspection of the shadow on the screen shows that light bends around the two sides of the wire enough to illuminate the *center* of its shadow. In addition, there will appear on the screen a number of narrow bright and dark bands parallel to the shadow of the wire. The explanation of these *diffraction fringes*, as they are called, is given in advanced texts. They are produced by a combination of interference and diffraction.

If an obstacle with a straight edge is put in the place of the wire *W*, the light will penetrate a noticeable distance into the geometrical shadow of the obstacle.

When sound-waves enter a window or other opening the width of which is comparable to their wave-length, they spread out in all directions from that opening. In a similar manner, when light passes through a very narrow opening, it spreads out. It is only when the width of the opening is comparable in size to the length of light-waves that the light spreads much. In Fig. 444 is shown an arrangement similar to that of Fig. 443. A narrow slit S_1 is illuminated by an intense source of light. A slit S_2 of adjustable width is placed between the slit S_1 and the screen. When the slit S_2 is made very narrow, the light passing through it spreads out over a considerable portion of the screen. This experiment shows an important fact: light passing through a very narrow slit spreads through a wide angle. This fact must be known before one can understand the principles involved in the diffraction grating, which are given in section 656.

---

* A projection lantern with the projection lens removed is suitable. The source is focused on the slit by the condenser lenses.

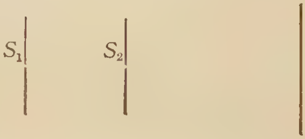


FIG. 444

An experiment similar to the one referred to in the last paragraph can be tried with very simple apparatus. As a source of light a single straight-filament street lamp is used. By means of a sharp knife a line is cut through the central part of a card, thus forming a narrow slit. By bending the card, this slit may be opened or closed as much as desired. The observer holds the card near his eye and looks through it at the source, opening and closing the slit, but keeping it parallel to the filament of the light source. He will not only see diffraction fringes, but he will also see that the light will spread out when it goes through a very narrow slit.

656. The diffraction grating. The *diffraction grating* is one of the most important optical instruments. With it much better spectra are obtained than with prisms, making it of great value in spectrum analysis. It is valuable for another reason: by its aid accurate measurements of wave-lengths can be made.

A transmission diffraction grating consists of a large number of fine, evenly spaced, and parallel slits. Originally these were made of fine wires, but better ones are made by ruling with a diamond a large number of parallel, equally spaced scratches on glass. The space between any two adjacent scratches is transparent and forms one of the slits. Reflection gratings are made by ruling fine parallel lines on a polished metallic surface. The space between the lines retains its reflecting power. Diffraction gratings are usually ruled with from 10,000 to 20,000 lines to the inch. The manufacture of accurate gratings is a very difficult piece of work and has been successfully done at only a few places. Copies, or replicas, of gratings have been made by pouring gelatin or some similar substance over a good grating and keeping it there until it solidifies. When the gelatin surface is stripped off, it retains an impression of the ruled surface, and forms a grating good enough for many purposes.

Before explaining the theory of the grating, a statement of the general character of the phenomena observed will be given. In Fig. 445, S is a slit which must be illuminated by a suitable source of white light, L_1 a lens which makes the light parallel, G the grating, and L_2 a lens for forming an image of the slit on the screen. On the screen is found, at the same place as though the grating G were not used, a white image, W, of the slit. This is called the *central image*. On each side of the central image is formed a spectrum with the red bent more than the violet, B_1R_1 in Fig. 445 (the figure shows only the light going to the white central image W and to the red image R_1 and the blue image B_1 on one side). Beyond each of the spectra B_1R_1 lies a second spectrum, B_2R_2, and, with some gratings, beyond these will be found one or more spectra. The spectra which lie next to and on each side of the central image are called the *first-order* spectra; the second set, the *second-order* spectra; and so on. When the grating is rather coarse, these spectra are crowded close together; but with gratings having a large number of openings to the inch, each of the spectra may have considerable length, greater than can be produced by a prism.

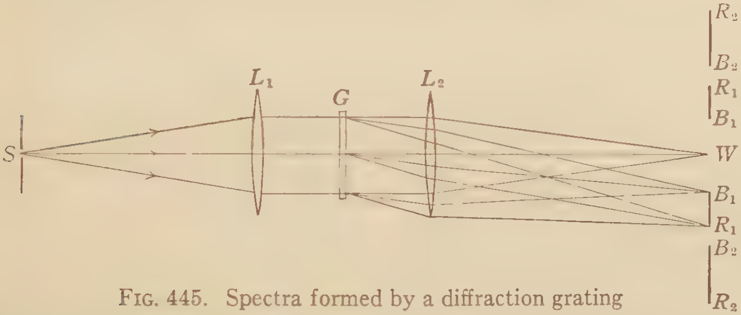


FIG. 445. Spectra formed by a diffraction grating

657. The theory of the diffraction grating. In a transmission grating (the only kind which will be explained here) the light passes through a large number of parallel slits. As stated in section 655, the light that passes through a narrow slit is diffracted,—it does not merely go straight through. In certain directions the light from one slit of a grating will interfere with that from some other slit. But there are other directions, as will be shown, in which the light from each slit is reënforced by light from all the others. We shall now find the directions in which waves from all the slits get "in step" and reënforce each other.

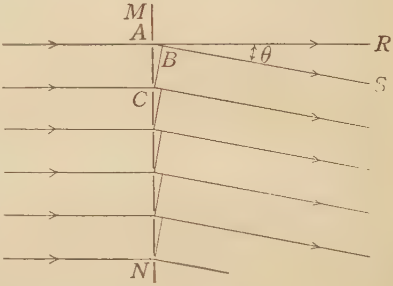


Fig. 446

Let the broken line *MN*, in Fig. 446, represent a greatly magnified portion of a grating. Parallel light is incident on the left-hand side. Some of this light will go directly through, and if a lens is used to focus it, a white central image will be formed. Let us find out what are the necessary conditions for light to travel parallel to the line *AS*. The line *CB* is drawn perpendicular to *AS*. When a lens is used to converge to a focus the parallel beam traveling in the direction *AS*, the light which emerges from the slit *A* will travel a distance *AB* farther in reaching that focus than the light from the second slit, *C*. Since the slits are equally spaced, there will be the same path-difference for any two adjacent slits. The path-difference between light emerging from any slit and the second slit from it will be twice the distance *AB*; from any slit and the third, three times; from any slit and the fourth, four times; and so on. *If the distance AB is 1 wave-length*, the waves from *all* the slits will reënforce each other at the focus; *hence that light which has a wave-length equal to AB will travel in a direction parallel to AS.*

It is obvious from the figure that the length of *AB* will depend on the size of the angle θ —the angle between the direct and the diffracted beam. When θ is small, AB is small; when θ is larger, AB is larger. It is always possible for some direction θ to exist for which AB will be equal to the wave-length of any one color, and in that direction the color will be reënforced. With a certain value of θ, AB will be equal to the wave-length of violet light, and violet will travel in that direction. Since red light has a longer wave-length than violet, the angle of diffraction, θ, must be larger for red than for violet. The other colors, which have wave-lengths between those of violet and red, will be diffracted in directions between these two. It follows from these considerations that a spectrum will be formed. While the figure shows the light traveling only on one side of the line AR, the same reasoning applies to directions on the other side of this line; hence a spectrum will be formed on the other side of the white central image. This is the way the two first-order spectra, which are indicated in Fig. 445, are formed.

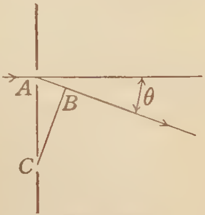


Fig. 447

The explanation of the second-order spectra is very similar to that of the first. The only difference is that the distance AB in Fig. 446 must be equal to 2 wave-lengths. This means that the angle θ must be larger, so that AB shall be twice as long as in the case of the first-order spectra. The third-order spectra occur when θ is so large that AB is 3 wave-lengths long.

To understand how wave-lengths of light are measured, it is necessary to find the relation between the distance AB and quantities which can be measured. The distance AC of Fig. 447 (which is a part of Fig. 446) is called the grating space. By the aid of a microscope the number of spaces on the grating to the inch or to the centimeter can be determined; from this the grating space is computed. Let this be called d, or $AC = d$. In the figure the angle ACB is, by construction, equal to θ. From the definition of the sine of an angle (sect. 590),

$$\sin ACB = \sin\theta = \frac{AB}{AC} = \frac{AB}{d},$$

or

$$AB = d\sin\theta.$$

For light-waves traveling in the direction making an angle θ with their original direction, AB equals 1 wave-length, say λ; hence

$$\lambda = d \sin \theta. \tag{28}$$

For the second-order spectra,

$$2 \lambda = d \sin \theta. \tag{29}$$

In an experiment, d and the angle θ are measured; then the wave-length λ can be computed from equation (28) or (29).

The following table gives the results of the measurement of a few wave-lengths:

Approximate limit of visible red, Fraunhofer line A		0.00007594 cm.
Red hydrogen line, Fraunhofer line C		0.00006563 cm.
Sodium lines	Fraunhofer line D_1	0.00005896 cm.
	Fraunhofer line D_2	0.00005890 cm.
Yellow mercury lines		0.00005791 cm. 0.00005770 cm.
Green mercury line		0.00005461 cm.
Blue hydrogen line, Fraunhofer line F		0.00004861 cm.
Approximate limit of visible violet, Fraunhofer line K		0.00003934 cm.

658. An alternative explanation of the grating. There is a relatively simple explanation of the diffraction grating based on Huygens' principle (sect. 570). Each of the slits of a diffraction grating may be regarded as a source of waves. When plane waves (in the last sections the discussion was limited to plane waves) fall on the grating, these sources will all be in phase. In Fig. 448 is shown a portion of a grating with wavelets emitted from each slit as a source. According to Huygens' principle the resultant wave is the surface which is tangent to the wavelets. In this figure the lines which are drawn tangent to the wavelets represent the waves which form the central image. But these are not the only tangent lines which can be drawn. Fig. 449 shows a line which is tangent to the first wave from slit 1, to the second from slit 2, to the third from slit 3, and so on. The position of such tangent lines can be traced out in Fig. 448, but for the sake of simplicity a group of them is shown in Fig. 449. These parallel lines represent the waves producing a first-order spectrum. The waves giving the second-order spectra are shown in Fig. 450. They are the lines formed by drawing a tangent to the second wavelet from slit 1, to the fourth from slit 2, to the sixth from slit 3, and so on.

659. X-ray spectra. Spectra of light-waves are produced by gratings in which the distance between the slits is of the order of magnitude of a wave-length of light. But X-rays are so much shorter than light-waves that no grating modeled after the one used with light-waves can be made. Furthermore, the absence of refraction of X-rays by prisms prevents the other common method of producing spectra. For a long time it seemed that no method would be devised for measuring their wave-lengths. But Laue and Bragg, by slightly different methods, showed that in the regular symmetrical arrangement of crystallographic units in crystals a natural diffraction grating was available (the crystallographic units are ordinarily atoms, but in organic crystals they are molecules).

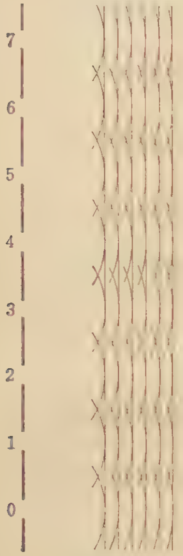


FIG. 448

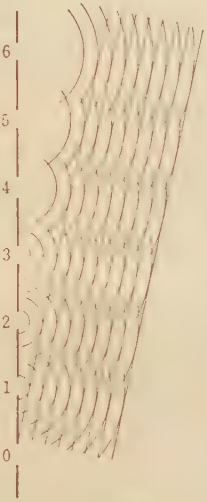


FIG. 449

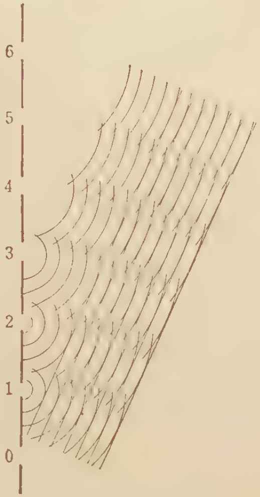


FIG. 450

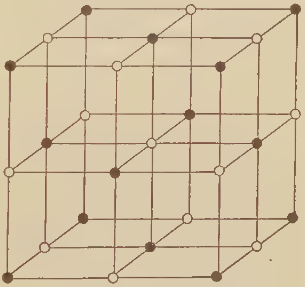


FIG. 451

In Fig. 451 is shown the arrangement of atoms in a simple cubic structure such as that possessed by crystals of common salt (sodium chloride), potassium chloride, potassium bromide, and potassium iodide. In the case of common salt there are equal numbers of sodium and chlorine atoms. The black circles in the figure represent sodium atoms, and the light circles chlorine atoms. In any structure such as this it is possible to find a number of planes, taken in different directions through the crystal, each of which is studded with atoms. The planes formed by the atoms are probably the most perfect that occur in nature. One of these will reflect X-rays, and the angle of incidence will always be equal to the angle of reflection. The reason for this equality can be understood from the following:

Let the line AC (Fig. 452) represent one of the crystal planes, and let the points represent, in an exaggerated way, atoms in this plane. When the incident wave AB strikes these atoms, each one tends to scatter the radiation in all directions. In the figure is shown the position of some of the wavelets at the time the wave would have reached DC. These wavelets are in phase along the line EC, and hence form a wave which travels in the direction indicated (this is another example of Huygens' principle). It can be shown that in other directions the waves from the atoms do not meet in phase, and hence tend to destroy each other by interference. The reflection, as may be seen from the figure, obeys the usual law of regular reflection, the angle of incidence being equal to the angle of reflection.

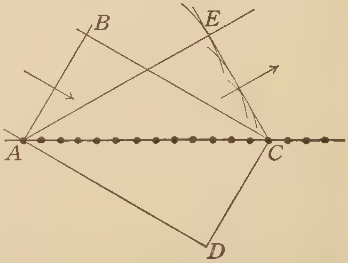


FIG. 452

If there were only one layer, or plane, containing atoms, as in Fig. 452, X-rays would be reflected from this plane at all angles of incidence; but there are vast numbers of parallel and equally spaced planes in any crystal of finite size. The reflection from these planes is somewhat similar to the reflection of light from the surfaces of thin films. The radiation reflected by one plane interferes with that from another except at certain angles of incidence. As shown in Fig. 453, when a beam of X-rays is reflected by two layers of atoms, the rays which are reflected from the second layer travel farther than those from the first, by a distance equal to $BC + CD$. If this distance is equal to 1, 2, or 3 wave-lengths, the two beams will reënforce each other.

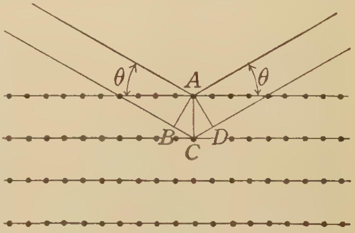


FIG. 453

From the definition of the sine of an angle (sect. 590),

$$\sin BAC = \frac{BC}{AC},$$

or

$$BC = AC \sin BAC.$$

The distance AC is the distance between the layers of atoms. We shall denote this distance by d. By construction the angle BAC is equal to the angle CAD and to θ, the angle of "grazing incidence." Then

$$BC = d \sin \theta.$$

But the path-difference of the two beams is $BC + CD$. Since $BC = CD$, the total path-difference is $2\,BC$. X-rays will be reflected, provided that

$$\lambda = BC + CD = 2\,BC.$$

Hence

$$\lambda = 2\,d \sin \theta. \tag{30}$$

X-rays will also be reflected when the path-difference is 2λ or 3λ.

The rays are reflected not only by two layers but by thousands of equally spaced ones. It can be seen that if the path-difference between the first and second layers is 1 wave-length, that between the first and third will be 2 wave-lengths, that between the first and fourth will be 3 wave-lengths, and so on; in other words, the rays reflected from all these layers will reënforce each other for the directions in which equation (30) is true.

The method of measuring the wave-length of X-rays consists of two steps: (1) to find d, the distance between the layers of atoms; (2) to measure the angle θ at which the rays are reflected.

The determination of the distance d is not so difficult as appears at first sight. The masses of atoms are now quite accurately known. Since we know the density of a given crystal (that is, the mass of 1 cubic centimeter), the number of atoms in a cubic centimeter is found by dividing the density by the mass of 1 atom. If the arrangement of the atoms is known, the distances apart can be computed from the number in a cubic centimeter. Fortunately, the atomic arrangement of some crystals is simple and known.

The value of d which has served as a standard for wave-length measurements has been either that of calcite or that of rock salt (between the planes known as "100"). For calcite, d is equal to 3.029×10^{-8} cm. For rock salt, d is equal to 2.814×10^{-8} cm.

The wave-lengths of X-rays, as measured by this method, range from 0.1×10^{-8} to 12×10^{-8} centimeters. The wave-lengths depend on what substance is used to radiate the X-rays: the elements having more massive atoms radiate shorter waves.

When the wave-length of the X-rays used is known, the distance between the planes in a crystal of unknown structure can be determined by the aid of equation (30). Hence, by reflecting X-rays from crystals, the most intricate systems of crystals are now being worked out. This method has furnished in recent years an entirely new and most fruitful method of working out the atomic arrangements in crystals.

CHAPTER XLVII

POLARIZATION OF LIGHT

Introduction, 660. Polarization of light by reflection, 661. Polarization by tourmaline, 662. An explanation of polarization, 663. Polarization of light proves that light-waves are transverse, 664. Double refraction, 665. Explanation of double refraction, 666. The Nicol prism, 667. Rotation of the plane of vibration, 668. The direction of vibration of plane-polarized light, and the electromagnetic theory, 669.

660. Introduction. It will be shown in this chapter that sometimes a beam of light possesses a property in one direction which it does not have in another. When this is true, the beam is said to be *polarized*. This term is used because the word *polarization* may be applied to anything which has a property in one direction that it does not have in another. A magnet tends to set itself in an approximately north-and-south direction. It has a property in one direction that it does not have in another; hence it is said to be polarized. When an electric current flows between two pieces of platinum which are immersed in acidulated water, a counter electromotive force is produced which acts in only one direction: it opposes the current flow. This is also a case of polarization.

The subject of polarization of light is a large and important branch of the science of optics; however, only a few of the simpler facts will be presented in this book.

661. Polarization of light by reflection. Ordinary light can be reflected by a piece of plate glass in any desired direction. But reflected light may partly lose this ability: there may be directions in which it cannot be reflected by a second piece of glass. Let light incident at an angle of about 57 degrees be reflected from a mirror M of unsilvered glass (Fig. 454). A second mirror, M', of unsilvered glass is set so that the angle of incidence is the same as that of the first mirror. In the positions shown in the figure both mirrors reflect light; but when M' is rotated (using the ray of light between M and M' as an axis) until the light should be reflected toward the student or away from him (Fig. 455), practically *no light will be reflected from the second mirror.* This may be stated in another way: If the angle of incidence for both mirrors M and M' is kept at about 57 degrees, the mirror M' will not reflect light when the plane of incidence[*] at M' is perpendicular to the plane of incidence at M. It will, however, reflect light in all other directions. In some manner the light has acquired such a property on reflection from M that it will not be reflected by M' in certain directions. The light is said to have been *polarized* by the reflection from M.

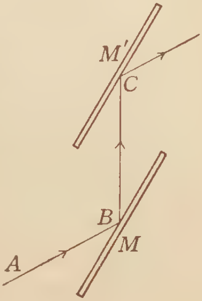


FIG. 454

The mirror M is called the *polarizer*, and the mirror M' (by means of which the polarization is detected) the *analyzer*. The angle of incidence (about 57 degrees for glass) at which this reflection must take place so that the light will be polarized is called the *angle of polarization.* Partial polarization will be produced at other angles of incidence.

Many other nonmetallic substances will polarize light by reflection. The magnitude of the angle of polarization is different for different substances.

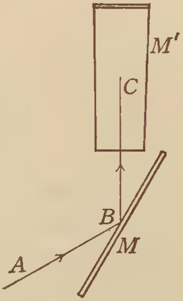


FIG. 455

662. Polarization by tourmaline. Light which passes through a plate of tourmaline[†] crystal becomes polarized. The polarization may be detected by using a second tourmaline. When the crystals are parallel, as in *ab*, Fig. 456, light passes through both. But when the crystals are partly crossed, as in *a'b'*, the light is dimmed; when they are set at right angles, as in *a''b''*, practically none of it goes through the two. Either crystal may be rotated in showing these effects. The crystal through which the light passes first is called the *polarizer*, and the second one is called the *analyzer*.

---

[*] The *plane of incidence* is the plane containing the incident ray and the normal to the reflecting surface.

[†] The plates of tourmaline referred to in this chapter are those which have been cut with their surfaces parallel to what is called the *optic axis* of the crystal.

Instead of two tourmalines, a piece of plate glass and one tourmaline may be used. If light is polarized by being reflected at the polarizing angle of glass, light will pass through the tourmaline when it is held in one position, but will not pass through it when it is turned through a right angle. Or if light is polarized by passing through a tourmaline, there will be two directions in which the transmitted light cannot be reflected by a piece of glass.

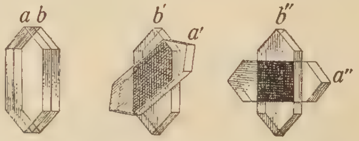


FIG. 456. Polarization of light by tourmaline

From the foregoing it is seen that no matter which method has been used for polarizing the light, the same results are obtained when tested by the reflection method. Or if tourmaline is used as the analyzer, the same results are obtained whether the light is polarized by reflection or by transmission through tourmaline. Hence it follows that light which has been reflected from glass at the polarizing angle acquires the same properties as light which has passed through a plate of tourmaline.

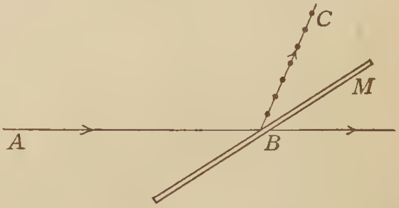


FIG. 457

663. An explanation of polarization. The experiments of the last two sections can be easily explained in case we assume that light-waves are *transverse*,—that they vibrate at right angles to the direction of their propagation. Ordinary light may vibrate in all directions at right angles to the direction in which it is traveling; but when it is polarized, the direction of the vibrations may be restricted. For example, when light is reflected from unsilvered glass at an angle of about 57 degrees, only those vibrations which are parallel to the plane of the glass surface are reflected. The incident ray (*AB*, Fig. 457) may contain light vibrating in all directions perpendicular to *AB*, but the reflected ray *BC* will contain only the vibrations which are perpendicular to the plane of the figure; that is, parallel to the surface of the mirror M. Because the vibrations are all in one plane, the reflected ray BC is said to be *plane-polarized.*

Since a plate of glass, set so that the angle of incidence is about 57 degrees, can reflect only those vibrations parallel to its plane, it will reflect no light if the incident light is vibrating as in the ray AB, Fig. 458. But if the light is vibrating perpendicularly to the plane of the figure (represented by the dots in the ray CD), part of it is reflected. If the mirror is turned so as to reflect light toward the student, keeping the angles of incidence approximately equal to 57 degrees, part of the ray AB would be reflected, but none of the ray CD.

A plate of tourmaline transmits only a part of the incident light, and that which it does transmit is vibrating in a direction parallel to the optic axis of the crystal. In other words, when ordinary light is incident on tourmaline, the transmitted light is plane-polarized. If the light is already plane-polarized, it may or may not go through a piece of tourmaline, depending on the direction in which the light is vibrating with respect to the plate. When the incident light is vibrating parallel to the optic axis of the crystal, it will go through; but when it is vibrating at right angles to this axis, it will not go through. Hence, as the crystal is rotated, in certain positions the light will go through, in other positions it will not.

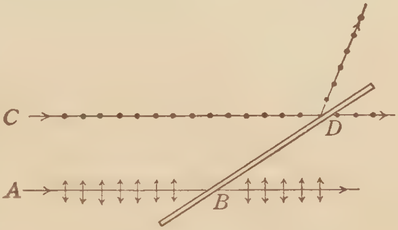


FIG. 458

In plane-polarized light the waves vibrate in only one plane.

664. Polarization of light proves that light-waves are transverse. As was explained in section 310, there are two types of wave-motion: *longitudinal* and *transverse.* In longitudinal waves the direction of vibration is the same as the direction of propagation; in the transverse type, vibrations are at right angles to the direction of propagation.

There is one distinct difference between the two types of waves: in one—the longitudinal—the vibrations are always symmetrical with respect to the direction of propagation, while in the case of the transverse type the vibrations are not symmetrical. Consider the case of longitudinal waves traveling up and down a long spiral spring hung from the ceiling. The longitudinal vibrations produce "compressions" and "rarefactions" which travel up and down the spring. These vibrations are completely symmetrical with respect to the direction the waves are traveling, for they will have exactly the same appearance when seen from any side. On the other hand, transverse waves sent along a rope do not have the same appearance when viewed from different positions. If the rope is horizontal and the vibrations are in a vertical plane, the appearance of the rope when seen from above or below is very different from the appearance when seen from one of the sides. Transverse waves are not symmetrical with respect to their direction of propagation. *Hence, whenever we show that a wave does not have complete symmetry about its direction of propagation, we know that it must be transverse.*

A simple illustration will give one test for symmetry. A certain plug fits in a hole. If this plug is rotated, and if it fits for all positions, then there is symmetry; but if the plug fits for one position and not for another, the plug and the hole do not have complete symmetry. Rotation is a test of symmetry. We do not need to know the details of why the plug does not fit: the mere fact that it does not is proof of a lack of symmetry.

Consider now the experiment of the rotation of a plate of tourmaline held in a beam of plane-polarized light. In one position it transmits the light, in another it does not. A rotation of the crystal is the only change. This experiment shows that *the light beam is not symmetrical.* The waves must therefore be transverse. Hence the phenomena of polarization of light afford conclusive evidence that light-waves are transverse.

665. Double refraction. When a piece of calcite is placed on a printed page, letters seen through it appear doubled. Each beam of light entering the crystal is broken up into two beams, as indicated in Fig. 459. This phenomenon is called *double refraction.*

In the case shown in Fig. 459 the two emitted beams are parallel and are separated only slightly; but when the calcite is cut into the form of a prism and backed by a glass prism, as shown in Fig. 460, the two emerging beams are diverging and are easily separated for use in experimental work. This combination is known as *double-image prism.*

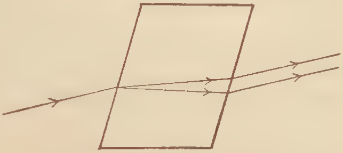


FIG. 459. Double refraction

When the two beams emitted by a double-image prism are tested by a tourmaline plate or are reflected by a piece of plate glass, it is found that *both beams are polarized.* It is also found that the directions of vibration of the waves in the two beams are in planes at right angles to each other, as indicated in Fig. 460.

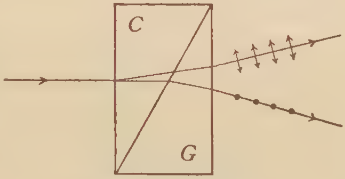


FIG. 460. Double-image prism

The intensity of the beam of light polarized by a tourmaline plate or by reflection from glass is small compared with the intensity of the original beam. The beams of polarized light transmitted by double-image prisms are relatively much more intense, and are hence better suited for experimental work in which polarized light is needed.

666. Explanation of double refraction. A block of wood has very different properties in different directions. Its thermal conductivity, its strength, and its elasticity are different in different directions. Let the broken line AA' (Fig. 461) represent the direction of the grain of a block of wood. Suppose that a transverse wave—some mechanical wave like a tremor—enters the front face of this block. Let us suppose, further, that the direction of vibration is parallel to the grain, in the direction BB'. This wave will travel through the wood with a certain velocity, which will depend, among other things, on the elasticity of the wood in the direction BB'. If the wave is vibrating across the grain, in the direction CC';

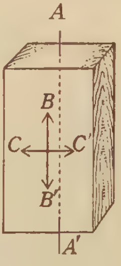


FIG. 461

the velocity of propagation will be quite different; for the elasticity of the wood in the direction CC' is different from that in the direction BB'. Hence the velocity of propagation of a transverse wave in wood will depend upon the direction of vibration of the wave.

Most crystals have different physical properties in different directions. As in the case of a block of wood, we should expect the velocity of propagation of a transverse light-wave to depend on the direction of vibration. Let the rectangle of Fig. 462 represent the face of a crystal. Let us suppose that when light, falling perpendicularly on this face, is vibrating in the direction BB', it travels through the crystal with its maximum velocity; but that when it is vibrating in the direction CC', it travels with the minimum velocity. If light vibrating in the direction AA' falls on the crystal, its vibrations will be broken up into two components—one parallel to BB', the other parallel to CC'. These two components will travel at different speeds in the crystal. *Because they travel at different speeds, they will be refracted different amounts.*

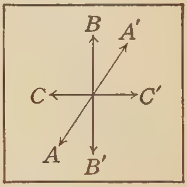


FIG. 462

In a similar manner, when ordinary light, which is vibrating in all directions, falls on a crystal face, it will be broken up into two beams—one vibrating parallel to BB', the other parallel to CC'. Since these beams travel at different speeds, there will be double refraction. The two beams will be plane-polarized, and the direction of vibration in one will be at right angles to that in the other.

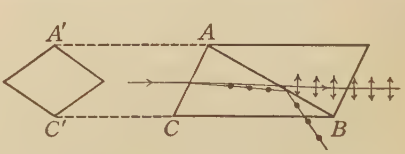


FIG. 463

667. The Nicol prism. The Nicol prism is one of the best forms of polarizing apparatus. Its use is so common that a brief description of its construction and use is given. A rhomb of Iceland spar, which is a pure, transparent form of calcite, is cut in halves along the line AB of Fig. 463; the cut faces are then polished, and cemented together again with Canada balsam. The shape of the end of the rhomb is shown by $A'C'$ (Fig. 463). Light entering one end is broken up into two beams by double refraction—one vibrating in the plane of the figure, the other at right angles to it. One of these beams is totally reflected by the layer of Canada balsam, and is usually absorbed by the case surrounding the rhomb. The other beam is transmitted. Hence a Nicol prism transmits plane-polarized light.

A Nicol prism may be used either as a polarizer or as an analyzer. When two Nicol prisms are placed in series, the first one acts as a polarizer and transmits plane-polarized light, and the second acts as the analyzer. If the second Nicol prism is set in the same relative position as the first, light is transmitted by it, as shown in Fig. 464. But if either prism is rotated through 90 degrees, the light will fail to pass through the analyzer. In Fig. 465 the polarizer *P* is shown rotated through 90 degrees. Rotation of *P* rotates the plane of vibration of the transmitted light. In this case light does not travel through the analyzer, but is totally reflected by the Canada balsam, as shown in the figure. The experiment of seeing light cut off by the rotation of one of the Nicol prisms is striking. Each of them is transparent to ordinary light; but if placed in series, the analyzer becomes opaque when the two prisms are "crossed."

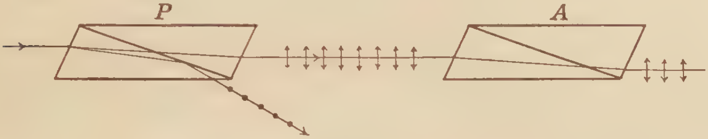


FIG. 464

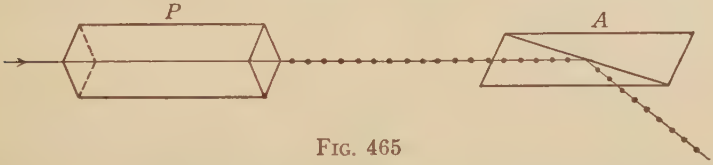


FIG. 465

The reason that the Canada balsam totally reflects one beam and transmits the other is a matter of some interest. In section 593 it was shown that light can be totally reflected only when entering a medium where its *speed* *is increased.* In the Canada balsam both beams have the same speed, since the balsam is not a crystalline substance. It so happens that the speed in Canada balsam is intermediate in value between the speeds which the two beams have in the spar. One of these beams, in going from the spar to the balsam, is *increasing* its speed; while the other, in going from the spar to the balsam, is *decreasing* its speed. The first of these will be totally reflected if the angle of incidence is greater than the critical angle. The prism is so shaped that the angle of incidence is always greater than the critical angle. But the second beam cannot be totally reflected at any angle.

668. Rotation of the plane of vibration. When plane-polarized light passes through a sugar solution, the plane of vibration of the light is rotated; hence the light which emerges from the solution is vibrating in a plane inclined at an angle to the plane of vibration of the incident beam. In Fig. 466 the light is polarized by the prism P, and is vibrating in the plane of the figure. While it passes through the sugar sirup at S, the plane of vibration is rotated—90 degrees in the case shown in the figure.

The magnitude of the rotation which is produced by a solution of sugar can be measured as follows: First, the analyzer is set for extinction; that is, so that no light passes through it when the sugar solution is not in the path of the light. After the solution is placed between the polarizer and the analyzer, the analyzer is rotated until the light is again extinguished. The angle through which the analyzer is rotated is the angle through which the solution rotates the plane of vibration.

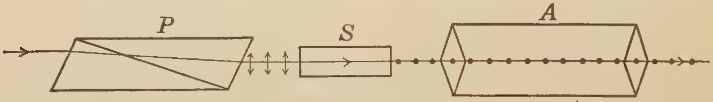


FIG. 466

The magnitude of the rotation by a solution depends on the concentration of the solution and its thickness. The measurement of the amount of rotation of the plane of vibration is a commercial test for the amount of sugar in any sirup. For example, it is possible to measure the amount of cane sugar in a mixture of cane sugar and glucose. The instruments used for testing the rotation in sugar solutions are called *saccharimeters.* Since different colors are rotated different amounts, monochromatic light must be used in these tests.

When plane-polarized light passes in certain directions through a quartz crystal, the plane of vibration is rotated. Turpentine, tartaric acid, and many other substances also rotate plane-polarized light. Not all active substances rotate the plane of vibration in the same direction. Most sugar solutions rotate the plane to the right, clockwise, as one looks through the solution toward the source of light; but levulose rotates it to the left.

669. The direction of vibration of plane-polarized light, and the electromagnetic theory. According to the now commonly accepted electromagnetic theory of light, a wave is complex, consisting of an electric field and a magnetic field (sect. 555). Both the electric and the magnetic components are vibrating at right angles to each other. In a plane-polarized wave of light the electric vibrations are all in one plane, and the magnetic vibrations in another at right angles to the plane of the electric vibrations. The direction of vibration which has been referred to in this chapter is the direction of the electric component. It would have made the explanations too complicated to introduce both components.

While the electromagnetic theory has had many triumphs, it must not be regarded as complete or final. Some recent results of the study of the spectra emitted by atoms throw much doubt on the sufficiency of this theory. It was pointed out in the last paragraph of section 563 that a satisfactory explanation of how energy is given up by an electron to a wave or of how it receives energy has not been found. There is another difficulty. The electromagnetic theory explains the wave-motion in terms of electric and magnetic fields. But we do not know what these fields are; there are serious objections to regarding them as disturbances in a fixed ether. However, a great deal of progress is being made. The advanced student will find that some wonderful work that throws much light on the structure of matter and on the way in which it radiates and absorbs energy is being done in the study of spectra, the ionization and disintegration of atoms, and other related fields.

ANSWERS TO PROBLEMS

Some of these answers are only approximations, for often the value of a quantity that is only approximately known must be used in the computations. For example, it may make a difference whether the specific gravity of mercury is taken as 13.6 or as 13.56.

CHAPTER I (PAGES 14–15)

1. 2.8 mi. northeast. **3.** 184 lb., 34 degrees north of east. **4.** 224 lb., $26\frac{1}{2}$ degrees from the 200-lb. force. **10.** 64.6 lb. **11.** 200 lb. **12.** 70.7 lb. **14.** 22 lb. **15.** 44 lb. **16.** 24.3 in. from the 10-lb. weight. **17.** 20 lb. **18.** 2.7 ft. from load. **19.** (*a*) 424 lb. (*b*) 424 lb. **20.** $1\frac{2}{3}$ ft. from heavier end.

CHAPTER III (PAGES 35–36)

2. (*b*) 9.84×10^5 bars; 14.31 lb./in.^{2} **3.** (*a*) 1.96×10^5 bars; 14.7 cm. mercury. (*b*) 0.617×10^5 bars; 4.64 cm. mercury. **5.** 0.09 lb./in.^{2} **6.** 134 lb. **7.** (*a*) 0.147×10^5 bars; (*b*) 2.14×10^5 bars; (*c*) 4.21×10^6 dynes. **9.** 1.48×10^{10} dynes. **12.** 112.7 gm. **13.** 3.33. **14.** 12,480 lb. **15.** $6\frac{1}{4}$ lb. **16.** 24.7%. **17.** (*b*) 205.9 lb. **18.** (*a*) 4.1; (*b*) 0.89. **19.** 6.66 lb.

CHAPTER IV (PAGES 51–53)

1. (*a*) 1.57×10^6 bars; 118.1 cm. mercury; (*b*) 1.14×10^6 bars; 85.5 cm. mercury. **2.** 718,000 dynes/cm.^{2} **3.** 2120 lb./ft.^{2} **4.** 964.600 bars. **5.** 7.9 oz. **7.** 78,000 cu. ft. **9.** 12.5 lb./in.^{2} **11.** 248 cu. m. **12.** 10.2 m. **13.** (*a*) 0.154 gm. (*b*) 0.095 gm. (*c*) + 0.059 gm. **14.** 5.74 m. **15.** 8.09 m. **16.** 1.03. **18.** 170 cm. mercury.

CHAPTER V (PAGES 64–65)

3. (*a*) 29.8 mi./hr. (*b*) 3.0 mi./hr. **6.** (*b*) 40 ft. **7.** (*a*) 7.5 ft. (*b*) 30 ft. **8.** (*b*) 125 cm. **10.** (*b*) 924 ft. **12.** (*b*) 100 ft. **13.** (*a*) 204 ft. (*b*) 116 ft./sec. **14.** (*a*) 189 ft./sec. (*b*) 5.90 sec. **15.** 465 ft. **16.** 5.8 ft. **17.** 153 ft.

CHAPTER VI (Pages 88–89)

1. (*b*) 64 cm. 2. (*b*) 10,000 dynes. 3. (*b*) 12,000 dynes. 6. Acceleration = 16 ft./sec.^{2} 7. 270 lb. 8. 880 sec. 9. (*a*) 6 ft./sec. (*b*) 25 ft. 10. 204.8 ft./sec. 11. (*b*) 108.9 cm./sec.^{2} 12. 131.2 lb. 13. (*a*) 4.69 slug-ft./sec. (*c*) 1 ft./sec. 14. (*a*) 8000 gm.-cm./sec. (*b*) 10 cm./sec. (*c*) 6000 dynes-sec. 15. (*b*) 9.88 ft./sec. 16. (*c*) 128 ft./sec. 17. 1.97×10^5 dynes. 18. 3.79×10^6 dynes. 19. 1815 lb. 20. 80.7 lb.

CHAPTER VII (Pages 107–108)

1. 74,900 ft.-lb. 2. 294×10^6 ergs. 3. 37.6¢. 5. 160,000 ft.-lb. 8. 424.26 ft.-lb. 9. (*a*) 90,000 ft.-lb. (*b*) 150 lb. 10. 1.58×10^8 ft.-lb. 11. 30.98 ft./sec. 12. (*a*) 22.5 ft./sec. (*b*) 25,312.5 ft.-lb. 13. (*a*) 100 cm./sec. 14. (*b*) 5×10^5 ergs. 15. (*a*) 1.16 slug-ft./sec. (*b*) 1743 ft.-lb. (*d*) 3.09 ft.-lb. 16. 5.71 H.P. 18. 0.606 H.P. 21. 16 H.P. 22. 1.09 H.P. 23. 312 lb. 24. 0.27 H.P. 25. 6250 lb. 26. 12 H.P. 27. 0.038 H.P.

CHAPTER VIII (Pages 118–119)

5. 2.45×10^6 dynes. 10. (*a*) 50 lb. (*b*) 200 lb. 11. (*a*) 66.7 lb. (*b*) 266.7 lb. 12. 420 lb. 13. 9048 lb.

CHAPTER IX (Pages 127–128)

3. (*a*) 20 lb. (*b*) 0.06. 4. 40,000 ft.-lb. 6. 150 lb. 7. 100 lb. 8. 9500 ft.-lb. 9. 100 lb. 10. 0.69 lb. 11. 125 lb. 12. 493.3 lb. 13. (*a*) 211.2 ft.-tons. (*b*) 0.64 H.P. 14. 393 ft.-lb./min.

CHAPTER X (Pages 137–138)

1. 188.5 cm./sec. 2. (*a*) 14.21×10^4 dynes. 3. 0.507 sec. 4. (*a*) 315.8 cm./sec.^{2} 5. 4.74×10^4 dynes. 6. 125.7 cm./sec. 7. 24.82 cm. 8. 2.007 sec. 9. 984 cm./sec.^{2} 10. (*a*) 1.0002 sec. (*b*) Lost 17.3 sec. 11. 3.55×10^6 ergs. 12. 1.26 sec.

CHAPTER XI (Pages 153–154)

1. 0.0391 slug-ft.^{2} 2. 12.34 ft.-lb. 3. 1.26 joules. 4. 9870 ft.-lb. 5. 1.12 slug-ft.^{2} 6. 2.83×10^5 dyne-cm. 7. (*a*) 3.51 slug-ft.^{2} (*b*) 6.6 lb.-ft. 8. (*a*) 14.06 slug-ft.^{2} (*b*) 17.66 lb.-ft.

CHAPTER XII (PAGES 161–162)

2. 0.00416 dyne. **3.** 1.003. **5.** 20.53 ft./sec.^{2} **6.** 25.8 lb. **7.** (*a*) 0.2131 dyne. (*b*) 0.0002174 gm.

CHAPTER XIII (PAGES 175–176)

1. 9 ft./sec. **2.** 39.6 lb./in.^{2} **3.** 17.7 ft./sec. **4.** 27.3 lb./in.^{2} **5.** 77 ft./sec. **6.** 32 ft./sec. **7.** 5568 ft.-lb. **8.** 1210 ft.-lb. **9.** 192 ft.-lb./sec. **10.** 3.8 ft.^{3}/min. **12.** 39,700 ft.^{3} **13.** 28,650 ft.^{3}/hr.

CHAPTER XIV (PAGE 193)

2. 1.54 cm. **3.** 0.40 cm. **4.** 0.40 cm. **6.** (*a*) 0.72 cm. (*b*) 0.37 cm.

CHAPTER XV (PAGE 202)

5. 0.0297 cm. **6.** 0.0283 in. **8.** 0.1 in. **9.** (*a*) 49×10^7 dynes/cm.^{2} (*b*) 4×10^{-4}. (*c*) 12×10^{11} dynes/cm.^{2} **10.** 0.0005. **11.** (*a*) 27.3 lb.-ft. (*b*) 857 ft.-lb./sec.

CHAPTER XVII (PAGES 226–227)

4. 0.030 cm. **6.** 400.43 cc. **7.** 6720 lb./in.^{2} **8.** 1124 cc. **10.** 497.85 gm. **11.** 5.4 cc. **12.** 93.0 lb./in.^{2} **13.** 39.9° C. **14.** 864.8 cc. **15.** 63.8 cm.

CHAPTER XVIII (PAGE 234)

5. 37.5° C. **6.** 17.5° C. **8.** 30.9° C. **9.** 13.4° C.

CHAPTER XIX (PAGES 248–249)

3. 344.5 cal. **4.** 26.4 m. **5.** 4.7 H.P. **6.** 5.1%. **7.** 1.41×10^4 lb. **8.** 1.9 H.P. **9.** 2.24 B.T.U. **10.** 427 gm. **11.** 1272.5 B.T.U.

CHAPTER XX (PAGES 264–265)

5. 1250 cal./sec. **6.** 20° C. **8.** 5.63×10^8 ergs/sec. **10.** 1550 kg.-cal.

CHAPTER XXI (PAGES 291–292)

3. 6.61 gm. **7.** 7.6° C. **8.** 161° C. **9.** 17.6° C. **11.** 240 gm. **13.** 31.2° C. **14.** 26.1° C. **15.** 15,026 cal. **16.** 1.51 kg.

CHAPTER XXII (PAGES 315–316)

1. 10.57 gm./m.^{3}; 62%; 1.05 cm. mercury. **2.** 3.4%. **7.** 11.86 gm.

CHAPTER XXIII (PAGE 324)

1. 36.4 H.P. **2.** 1.2 lb. **3.** 11.7%. **4.** 1.57 H.P.-hr.

CHAPTER XXIV (PAGE 359)

1. 12.4 ft./sec. **2.** 5.7 ft./sec. **4.** (*a*) 26.4 cm. (*b*) 571 cm. **5.** 5626 ft. **7.** 278 sec.$^{-1}$. **8.** 1.176. **11.** 2906 sec.$^{-1}$.

CHAPTER XXV (PAGE 378)

2. 40 C.G.S. units. **3.** (*a*) 0.5 C.G.S. unit. (*b*) 0.1 C.G.S. unit. **4.** 0.225 C.G.S. unit. **6.** 0.89 C.G.S. unit. **8.** 0.6 C.G.S. unit. **9.** (*a*) 240 dynes. (*b*) 4800 dynes-cm.

CHAPTER XXVI (PAGE 406)

2. 0.6 C.G.S. unit. **5.** (*a*) 0.064 C.G.S. unit. (*b*) 0.209 C.G.S. unit. **6.** (*a*) 0.4 C.G.S. unit.

CHAPTER XXVIII (PAGES 433–434)

1. 5.3 amp. **3.** 80 volts. **4.** (*b*) 1080 volts. **5.** (*b*) 103,680 joules. **6.** 103 volts. **7.** (*a*) 3.2 volts. **10.** 500 watts. **11.** 0.6¢. **12.** 8.8¢. **13.** 3.2¢. **14.** (*a*) 28 volts. (*b*) 5 kw. (*c*) 1120 watts.

CHAPTER XXIX (PAGES 442–443)

2. 9331 cal.* **3.** 667 sec.* **5.** (*b*) 216,000 cal.* **7.** 58.4° C.* **9.** 22¢. **10.** (*a*) 125 watts. (*b*) 12.5 kw. (Power factor = 1.)

CHAPTER XXX (PAGES 460–463)

1. 4.33 ohms. **2.** (*a*) 2.5 amp. (*b*) 275 watts. **3.** In 5-ohm, 0.36 amp.; in 10-ohm, 0.18 amp. **4.** (*a*) 112.2 volts. (*b*) 0.22 ohm. (*c*) 22 watts. **5.** (*a*) Loss in battery, 8 volts; in wires, 3.2 volts; in instrument, 8.8 volts.

---

* Equation (35) was used.

6. (*a*) 0.3 amp. **7.** 0.353 amp. **8.** 0.273 amp. **9.** (*b*) 1.0 volt. **10.** 32 volts. **11.** (*a*) 0.5 amp. (*b*) Dry cell, 1.4 volts; gravity cell, 0 volt. **12.** (*a*) 0.1 amp. (*b*) Dry cell, 1.48 volts; gravity cell, 1.2 volts. **13.** (*a*) 135 volts. (*b*) 600 watts. **14.** (*a*) 5 amp. (*b*) 30 volts. **16.** Shunt, 19.98 amp.; coil, 0.02 amp. **17.** 0.00101 ohm. **18.** 33.5 volts. **19.** (*a*) 40 amp. (*b*) 200 volts. **20.** 195 volts. **21.** 164 volts. **22.** (*a*) 3.46 ohms. (*b*) 2.7 volts. **23.** (*a*) 10 amp. (*b*) 116 volts. (*c*) 110 volts. **24.** (*a*) 600 watts. (*b*) 450 watts. (*c*) 105 volts. **25.** 0.2 ohm.

CHAPTER XXXI (PAGE 476)

1. 1.184 gm. **2.** 0.745 amp. **3.** 0.06276 gm. **4.** (*a*) 0.4 amp. (*b*) 0.00502 gm. **5.** 1.664×10^{-24} gm.

CHAPTER XXXII (PAGES 493–494)

1. 41.9 dynes. **2.** 16 dynes. **6.** (*b*) 270 dynes-cm. **7.** (*b*) 120 dynes. **8.** 0.008008 ohm. **9.** 267.7 ohms. **10.** 9086 ohms.

CHAPTER XXXIII (PAGES 521–522)

2. 0.4 volt. **3.** 5 volts. **4.** 0.0138 volt. **6.** 0.005 volt. **7.** 500 cm./sec. **8.** 0.00210 volt. **9.** 200,000. **10.** 0.013 sec. **12.** 320 volts. **14.** 180 R. P. M.

CHAPTER XXXIV (PAGE 530)

5. (*a*) 2513 lines.

CHAPTER XXXV (PAGES 548–549)

1. (*b*) 0.000033 m.f. **2.** 710 m.f. (approx.). **3.** (*a*) 112.5 C.G.S.E.S. units. (*b*) 0.000125 m.f. **6.** 0.0119 m.f. **7.** (*a*) 0.286 m.f. **9.** 7.2 joules. **11.** 133.3 volts.

CHAPTER XXXIX (PAGE 591)

1. 91,888 sec.^{-1} **2.** 0.0253 henry. **3.** 7.03 m.f. **4.** 300,000 sec.^{-1} **5.** 750,000 sec.^{-1}

CHAPTER XL (PAGE 605)

1. 2.525 foot-candles. **2.** 12.99 ft. **4.** 50.4 watts. **5.** 10.67 candles.

CHAPTER XLI (Page 622)

1. (*a*) + 20 cm. (*b*) + 60 cm. (*c*) − 30 cm. **3.** (*a*) − 8 cm. (*b*) − 5 cm. **5.** (*a*) − 10 cm. (*b*) 8 cm. **6.** (*a*) − 8.57 cm. (*b*) 4.28 cm. **7.** + 5.83 in. **8.** 20 in.

CHAPTER XLII (Page 638)

1. 1.33. **2.** 0.376. **3.** 0.855. **6.** 0.443.

CHAPTER XLIII (Pages 651–652)

1. (*a*) + 60 cm. (*b*) + 120 cm. (*c*) − 60 cm. **2.** + 10 cm. **4.** + 24 cm. **5.** + 40 cm. **6.** (*a*) − 13.3 cm. (*b*) − 10 cm. **8.** + 16.7 cm. **9.** 8 ft. 9 in. **10.** + 26.7 cm. **11.** − 10 cm. **12.** − 40 cm. from second lens. **13.** + 10 cm. from second lens. **15.** (*b*) + 133 cm. **17.** + 24 cm.

CHAPTER XLIV (Page 664)

1. 0.29 in. **2.** 9.73 in. **3.** 29. **4.** 9.33. **5.** 390. **6.** 355. **7.** 479. **9.** (*a*) − 62 in. from eyepiece. **10.** 0.43 in.

INDEX

Aberration, of light, 602; spherical, 621, 648; chromatic, 648
Absorption spectra, 669
Acceleration, defined, 54; produced by forces, 66–86; angular, 141; table of numerical values of g, 159
Accumulator, electric, 474
Achromatic refraction, 649
Adiabatic compression and expansion, 241
Air, weight, 37; pressure, 43; buoyancy, 45; liquefaction, 289
Air columns, vibrations of, 355–357
Air pumps, 49
Alpha radiation, 571–576
Ammeter, 415, 487, 488
Ampere, unit of current, 408, 552, 553; relation of, to coulomb, 417; international, 471, 553; defined, 479
Amplifiers, vacuum-tube, 565–567
Amplitude defined, 133
Analyzer, 699
Anastigmatic lens, 650
Angle of repose, 124
Angular acceleration, 141
Angular momentum, 150
Angular speed, 140, 141
Angular velocity, 140–146
Anticyclone, 309–310
Arago, 506
Archimedes' principle, 28, 45
Aristotle, 46, 597
Armature of a motor, 490
Aspirator, 167
Astigmatism, 650
Aston, 579
Atmospheric refraction of light, 636
Atom, mass, 178–179, 470, 569; size, 178, 578; electrical, 397, 568; number of atoms in one cubic centimeter, 470; disintegration, 573–576; structure, 577
Atomic number, 578
Atomic theory of matter, 177–181, 577
Atwood's machine, 74

Balance, 115
Balance-wheel compensation, 214
Banks, Joseph, 465
Bar defined, 24, 26
Barometer, mercury, 41; aneroid, 42
Baseball, curving, 166
Beats, 347
Becquerel, 570
Bending of beam, 201
Bernoulli's theorem, 168
Beta radiation, 572–576
Bicycle rider, 85
Binoculars, 662
Biprism, Fresnel, 683
Boiling, 278–280, 281, 283
Boiling point, effect of pressure upon, 278; effect of elevation upon, 279; table for water, 280; table for different liquids, 281; of solutions, 283
Borda, 17
Bouguer, 155
Bourdon spring gauge, 45
Boyle's law, 47
Boys, C. V., 156, 159
Bradley, 602
Bragg, 695
Bright-line spectra, 668, 671
British Absolute system, 73
British Engineering system, fundamental units, 16; compared with C.G.S. system, 105
British thermal unit defined, 229
Browne, Sir Thomas, 401
Bulk modulus, 199; table of values, 200

Caloric, 228, 235
Calorie defined, 229
Calorimetry, 230
Camera, pinhole, 592; photographic, 653
Candle, international, 595
Candle power, 595
Canton, John, 34, 402

Capacitance, 537–545; defined, 539; dependence on medium, 540; units, 541, 553
Capacity, thermal, defined, 231; electrical, *see* Capacitance
Capillarity, 186–188
Carbon-dioxide snow, 286
Carlisle, Sir Anthony, 410, 465
Carnot engine, 322
Cartesian diver, 30
Cathode dark space, 557
Cathode rays, 558–561
Cavendish, 403, 541
Cell, voltaic, 407, 465; in series, 454; in parallel, 454; in series-parallel, 455; theory of, 471; gravity, 473; dry, 474; storage, 474–476
Center of gravity, 12
Central force, 80
Centrifugal drier, 85
Centrifugal force, 81–86
Centripetal force, 79–86
C.G.S. system, fundamental units, 16; compared with B.E. system, 105; electrostatic and electromagnetic systems, 551–553
Charges of electricity, 379–401, 531–539; from dry cells, 391; from power circuit, 392; unit, 396; law of force, 396; natural unit, 397, 568; quantitative relation to voltage of a condenser, 537
Chemical effects of electric currents, 409, 464–476
Chromatic aberration, 648
Cloud, 297, 301
Coefficient, starting friction, 120; sliding friction, 121; rolling friction, 124; elasticity, *see* Modulus; linear expansion, 212; volume expansion, 215; thermal conductivity, 253; electrical resistance, 445; self-induction, 512
Coil, induction, 507
Colors, spectrum, 634, 635, 665–671; mixing, 677; complementary, 677; mixing of pigments, 678; thin films, 684
Columbus, 366
Combustion, heats of, 245; numerical values, 246
Commutator, 490
Compass, magnetic, 362
Components of a vector, 7
Compressibility of liquids, 34; of gases, 47
Compression of gases, 240; adiabatic, 241
Condenser, electrical, 401, 403, 535–546; in parallel and series, 542; energy of, 545
Conduction, of heat, 250–254; of electricity, 381; of electricity through liquids, 464–476; of electricity through gases, 555–569
Conductivity, thermal, 253; numerical values, 254
Conductors of electricity, 381, 423
Conjugate foci, with mirrors, 612; with lenses, 642
Conservation, of momentum, 78; of energy, 103; of angular momentum, 151
Convection, 254–257, 302–303. *See also* Winds
Cookers, fireless, 252; electric, 436
Cooking of food, 252
Cooling, rate of, 263
Coronas, 676
Corpuscular theory of light, 598, 605, 624, 625
Coulomb, 373, 396, 403; unit of charge, 397; relation of, to ampere, 417; relation of, to electromagnetic unit, 553
Coulometer, 471
Counter electromotive force, 457; of motor, 515
Couple defined, 13
Cream separator, 85
Critical angle, 629; numerical values, 632
Critical pressure, 287–288
Critical temperatures, 287
Crookes, Sir William, 561
Crookes dark space, 557
Cruickshank, 465
Crystals, used in measurement of X-rays, 694–697; distance between crystallographic units, 697
Cunæus, 403
Curie, 570
Current of electricity, unit of, 408, 471, 478, 551, 552, 553; relation to quantity, 416
Current-magnetic-field force, 413, 477, 480–491; relation to induced electromotive force, 505
Curve of a baseball, 166
Cyclones, 307, 310
Cymometer, 38

Dalibard, 402
Dalton, 282
Damped electrical oscillations, 582
Dark-line spectra, 669–671
Davy, Sir Humphry, 236, 465
Debierne, 571
Declination, magnetic, 366
Degradation of energy, 247
Density, defined, 30; table of numerical values, 32; table for gases, 47
Dew, 293
Dew point, 294
Dewar flask, 261
Diamagnetic substances, 528
Dielectric constant, 540; table, 541
Diesel engine, 324
Differences of potential, 418–421, 424, 531–534; quantitative definition, 425; measurement, 427; quantitative relation to charge on condenser, 537
Differential expansion, 211
Differential pulley, 113
Diffraction, sound-waves, 688; light-waves, 688–694; X-rays, 694–697
Diffraction grating, 690–694
Diffusion, liquids, 190; solids, 190; gases, 190
Dioptric, or diopter, defined, 647
Dip, magnetic, 367
Discharge, oscillating electrical, 581–583
Dispersion, 634, 649
Dissipation of energy, 247
Distillation, 282; fractional, 283
Distortion of images, 651
Distribution of electric charges, 393
Divided circuits, law of, 450
Doorbell, 412
Doppler's principle, 348, 672
Double refraction of light, 702–704
Double-image prism, 703
Drier, centrifugal, 85
Du Fay, 402
Dulong, 217
Dynamo, direct-current, 513; alternating-current, 515, 585
Dyne, relation to pound, 1; defined, 72

Earth, mass, 158; mean temperature, 259; a magnet, 365–368
Echo, 338
Edison storage cell, 476
Efficiency, of machines, defined, 118; of engine, 320, 321, 322, 324
Elasticity, 194–201
Electricity, 379–591
Electrification, 380
Electrochemical equivalent, 468
Electrochemical equivalents, table, 469

Electrolysis, 409, 464–476; Faraday's laws, 468
Electromagnetic system of units, 551–553
Electromagnetic theory of light, 599, 707
Electromagnetic waves, 587–591
Electromotive force, 408; defined, 431, 453; counter, 457, 515; induced, 495–521
Electron, 178; in theory of electricity, 383; emission of electrons, 561, 564–568, 572, 574–576; mass, 569; in atom, 577; size, 578
Electrophorus, 390
Electroplating, 410, 466–468
Electroscope, 384, 388, 389
Electrostatic machines, 391
Electrostatic system of units, 551, 552, 553
Energy, 95–100; table of numerical equivalents of units, 92, 239; potential, 97, 100; kinetic, 97, 98, 148; hydraulic, 170–173; in magnetic field, 376, 513; transmission of electric, 440, 518; of electrical condenser, 545
Engine, steam, 317–322; heat, 317–324; Carnot, 322; gas, 323; Diesel, 324
Equilibrium, first and second conditions, 11
Erg defined, 92
Evaporation, 179, 273–290
Exchanges, Prevost's law, 258
Expansion, thermal, 210–226; differential, 211; coefficient of linear, 212; coefficient of volume, 215; relative, 216; absolute, of a liquid, 217; of gases, 219–225, 240–245; adiabatic, 241; free, 243
Eye, 655

Falling bodies, laws, 59; acceleration, 60, 73, 159
Farad, unit of capacitance, 541, 553
Faraday, ice-pail experiment, 399; laws of electrolysis, 468; induced electromotive force, 495, 497, 499, 506; dielectrics, 541
Field, electric, 397

Field, magnetic. *See* Magnetic field
Films, thin, color of, 684
Filter pump, 167
Fireless cookers, 252
Fluid friction, 125, 163
Fluid theories of electricity, 382
Fluids, in motion, 163;175. *See* Gases and Liquids
Fluorescence, 678
Flux, magnetic, 526
Focus, principal, of mirror, 611, 618; of lens, 640
Fog, 297
Food energy needed by a man, 98, 245
Foot-candle, 595
Foot-pound, 91, 92
Foot-ton, 92
Force, 1–14; parallelogram of forces, 2; polygon of forces, 3; equilibrium of forces, 4; components, 7; moment of, 9; units, 18, 71; produces acceleration, 66–86; central, or centripetal, 79–86; centrifugal, 81–86; involved in thermal expansion, 219; between magnetic poles, 373; between electric charges, 396, 541; between current and magnetic field, 413, 477–491
Foucault, 147, 603, 625
Franklin, Benjamin, 402, 403, 404
Fraunhofer lines in spectrum, 670
Freezing. *See* Fusion
Freezing point. *See* Melting point
Fresnel, 682, 683
Friction, 120–127; starting, 120; sliding, 121; tables of coefficients, 122; rolling, 124; fluid, 125, 163
Frost, 260, 293
Fuses, electric, 436
Fusion, 266–273; heat of, 268

Galileo, 129, 600
Galvanometer, 414, 486; ballistic, 524
Gamma radiation, 572–577
Gas engine, 323
Gas law, general, 221
Gases, 37–51; kinetic theory, 39; molecular theory, 39; specific gravity, 46; density, 46; expansion, 219–225, 240–245; general law, 221; table of specific heats, 245; conduction of electricity through, 555–569
Gasoline, electrification, 281, 380
Gauges, pressure, 44
Geissler tubes, 558
Generator, direct-current, 513; alternating-current, 515, 585; vacuum-tube, 586
Gilbert, William, 365, 401
Gradient, temperature, 254
Gram, 18
Gram-calorie defined, 229
Grating, diffraction, 690–694
Gravitation, 155–161; law, 156; constant, 156, 159
Gravity, center of, 12; acceleration caused by, 60, 73; variations in acceleration, 159
Gray, Stephen, 402
Guericke, Otto von, 39, 402
Gyration, radius, 150
Gyroscope, 144
Gyroscopic motion, 141

Halos, 676
Harmonics, 344; in strings, 350–352; in air columns, 356
Harmonious combinations of sounds, 358
Head of a liquid, 25
Heat, 203–324; units, 229; mechanical equivalent, 239
Heat of combustion, 245; numerical values, 246
Heat of fusion, defined, 268; numerical values, 269
Heat of vaporization defined, 284
Heater, electrical, 435
Heating, of houses, 255; electrical, 409, 435–441
Henry, Joseph, 495, 497
Henry, unit of inductance, 512, 553
Hero, steam engine of, 317
Hertz, 589
Hooke's law, 194
Horse power defined, 101
Horse-power hour, 92, 103
Hull, 599
Humidity, absolute, 295; relative, 295; table of absolute, 296
Huygens' principle, 607
Hydraulic press, 27, 117
Hydraulic ram, 174
Hydrometer, 34
Hydrostatic bellows, 27
Hydrostatic press, 27, 117
Hygrometers, 297
Hysteresis, 528

Ice machine, 288–289
Ice-pail experiment, 399

Illumination, intensity, 593–597
Images, real and virtual, defined, 610
Impulse defined, 76
Inclination, magnetic, 367
Inclined plane, 93, 114
Index of refraction, defined, 627, 629; relative, 633; absolute, 633
Induced electromotive force, 495–521
Inductance, 512
Induction, magnetic, 368–370, 526; electrostatic, 386; electromagnetic, 495–521
Induction coil, 507
Inertia, 67; moment of, 148; in electrical circuit, 510, 581
Infra-red light, 673, 674
Insulators, 381
Intensity, of sound, 343; of magnetic field, 374; of electric field, 397; of magnetic field produced by current, 478–479; of illumination, 593–597
Interference, of sound-waves, 346; of light-waves, 680–687
Interlinkage defined, 512
Internal-combustion engine, 323
International electrical units, 553
Interval, musical, 357–358
Ionization, of liquids, 464; of gases, 555, 571, 572
Ions, in electrolytes, 466; in gases, 555
Iron, magnetization, 368, 523–529
Isogonic lines of United States, 367; map, 366
Isotope, 178, 578

Jackscrew, 115
Joule, determination of mechanical equivalent of heat, 237; law of electric heating, 437, 439
Joule, unit of work, defined, 92
Joule-Thomson effect, 244

Kelvin temperature scale, 204, 222
Kelvin wave-apparatus, 326
Kilogram, standard, 18
Kilogram-calorie defined, 229
Kilowatt defined, 102
Kilowatt-hour, 92, 103, 430
Kinetic energy, 97–99; of rotating bodies, 148
Kinetic theory, of gases, 39, 241; of evaporation, 274
Kinnersley, 404
Kleist, Von, 403
Kundt's tube, 354

Lamp, incandescent, 436
Land breeze, 256
Langmuir, 437
Lantern, projection, 655
Laue, 695
Lebedew, 599
Lenses, 639–652
Lenz's law, 503
Lever, 10, 109
Leyden jar, 400, 403, 535
Lifting magnet, 413
Light, 592–707
Lightning, 313, 389, 402
Lightning rods, 394, 395, 402
Lines, magnetic, 364, 375; electrical, 398
Liquefaction of air, 289
Liquids, 20–36; in motion, 163–175
Lodestone, 360
Longitudinal waves, 326, 327
Luminescence, 671, 679

Machines, simple, 109–119
Magdeburg hemispheres, 39
Magnetic circuit, 529
Magnetic field, 363; strength, or intensity, 374; represented by lines 375; energy, 376, 513; produced by electric current, 410, 478; used to define unit current, 478; rotating, 517
Magnetic induction, 368, 526
Magnetic permeability, 528
Magnetism, 360–376; produced by electric current, 410; of iron, 523–529
Magnetizing force, 525
Magnification, law of, in mirrors, 615; in lenses, 647; magnifying glass, 657; telescope, 660
Magnifying glass, 656, 657
Map of United States, magnetic, 366
Mariotte, 47
Mars, mean temperature, 259
Mass, units, 18; distinguished from weight, 69, 73, 74
Maximum density of water, 218
Maxwell, 541
Mechanical advantage defined, 110
Mechanical equivalent of heat, defined, 239; numerical values, 239; determination by electrical method, 438
Melting. *See* Fusion
Melting point, numerical values, 267; effect of pressure on, 270–272; of solutions, 273

Meteorology, 293–315
Meter, standard, 17
Meter-candle, 596
Michelson, 605
Microfarad, 541
Micromicrofarad, 541
Micron, 18
Microscope, simple, 656; compound, 658
Millikan, R. A., 397, 470, 568
Minimum deviation by prism, 667
Mirage, 637
Mirrors, plane, 609; concave, 611; convex, 618; parabolic, 621; Fresnel, 682
Mixtures, method of, 230
Modulus, Young's, 196; numerical values of Young's, 197; of volume elasticity, 199; numerical values of bulk, 200; rigidity, 200; numerical value of rigidity, 201
Molecular forces, 180–189
Molecular theory, of gases, 39, 241; of matter, 177–181; of heat conduction, 250; of magnetism, 371
Molecules, number of, 179, 470
Moment, of force, defined, 9; of inertia, 148; of magnet, 376
Momentum, defined, 75; conservation of, 78; angular, 150
Moon, and law of gravitation, 157
Motion, uniformly accelerated, 54–65; general laws, 56–57; effects of forces on, 66–86; Newton's laws, *see* Newton; simple harmonic, 129–133; rotary, 139–153
Motor, direct-current, 489–491; counter electromotive force, 515; induction, 517
Musschenbroek, 403

Natural unit of charge, 397, 568
Newcomb, 605
Newton, first law of motion, 66, 77; second law of motion, 77; third law of motion, 78; law of gravitation, 156; law of cooling, 263; corpuscular theory of light, 598; dispersion of light, 634–636, 665; rings, 687
Nichols, Ernest F., 599
Nicholson, William, 410, 465
Nicol prism, 704
Nodal points in lens, 641
Nucleus of atom, 178, 383, 577

Ocean currents, 256
Octave, 358
Oersted, 410
Ohm, unit of resistance, 408, 422, 552, 553
Ohm's law, 409, 422, 451–458, 511; general form, 455
Opera glass, 660
Optical center of lens, 641
Optical instruments, 653–664
Organ pipe, 356
Oscillations, high-frequency electrical, 580–591
Osmosis, 192
Osmotic pressure, 192
Overtones, 343; in strings, 350–352; in air columns, 356

Paradox, hydrostatic, 25
Parallel resistances, law of, 448
Parallelogram of forces, 2
Parhelia, 676
Pascal's principle, 27
Pendulum, simple, 134; physical, or compound, 136; torsion, 136; Foucault, 147; compensated, 213
Period of vibratory motion, 133
Permeability, magnetic, 528
Perpetual motion, 103
Petit, 217
Phase of vibratory motion defined, 133
Phosphorescence, 678
Photo-electric effect, 568
Photographic camera, 653
Photometry, 595
Pigments, mixing, 678
Pinhole camera, 592
Pitch, musical, 341; effect of motion on, 348
Pitot tube, 175
Plane, inclined, 93, 114
Plato, 597
Points, electrical action of, 394
Polarization, of voltaic cells, 473; of light, 698–707
Polarizer, 699
Poles, magnetic, 361–366; law of force between, 373; unit, 374
Potential, electrical, 417–421, 424, 451, 531–534; quantitative definition, 425
Potential energy, 97, 100
Pound, relation to a dyne, 1; unit of force, 18, 72

Poundal, 73
Power defined, 101; in an electric circuit, 428, 431
Power of a lens, 647
Poynting, 156
Practical system of electrical units, 552
Precession, 144
Pressure, defined, 21; in liquids, 21–27; units, 26; table of value of units, 26; transmission of, 27; in gases, 37–44; gauges, 44; changes when speed of liquid changes, 163–169; effect on melting point, 270; of saturated vapor, 275–289; effect on boiling point, 278; of light, 599
Pressure-depth formula, 24
Prevost's law of exchanges, 258
Principal focus, of mirror, 611, 618; of lens, 640
Prism, refraction of light, 634, 665; penta, 663; double-image, 703; Nicol, 704
Prism binocular, 661
Projectiles, 61
Projection lantern, 655
Pulley, 112–114
Pump, lift, 48; force, 49; centrifugal, 49; air, 49; compression, 50; filter, 167
Pythagoras, 129

Quality of sound, 343, 352
Quantity of electricity, unit, 397, 551, 552, 553; relation to current, 416; relation to voltage of a condenser, 537
Quantity of heat, 228–233
Quantum theory, 600

Radian defined, 82
Radians per second or minute, and linear speed, 140, 141
Radiation, thermal, 257–263
Radioactivity, 570–579
Radium, 570–579
Rain, 301–304
Rainbow, 674
Raindrops, speed of fall, 126, 298
Rainfall, averaging of, 315
Ram, hydraulic, 174
Range-finder, 663
Real image defined, 610
Rectification of alternating current, 564
Rectifier, thermionic, 564
Rectilinear lens, 651
Reflection, of heat, 258; of sound, 338; of light, 606–622; of X-rays, 696
Refraction, of sound, 339; of light, 623–638; double, 702–704
Refrigerating machine, 288
Regelation, 271
Regnault, 34, 217
Repose, angle of, 124
Resistance, electrical, 408, 422, 444–451; of different materials, 423; temperature coefficient, 445
Resistance box, 445
Resonance, 345; electrical, 584
Return shock, 389
Rheostats, 445
Richmann, 402
Right-hand rule, rotational vectors, 142; magnetic field, 411
Rigidity modulus, 200
Rings, Newton's, 687
Ritter, 410
Roberts-Austen, 190
Rods, vibrations of, 353
Rolling friction, 124
Römer, 600
Röntgen rays. *See* X-rays
Rotary motion, 139–153
Rotating magnetic field, 517
Rotation of earth, effect on projectiles, 147; effect on winds, 147, 306; shown by Foucault's pendulum, 147
Rotation of plane of vibration of light, 706
Rumford, Count, 235
Rutherford, 577

Saccharimeter, 706
Sailboat, 8
Scale, major, 358
Schmidt, G. C., 570
Screening, magnetic, 361, 369; electrical, 393
Screw, 115
Sea breeze, 256
Self-induction, 509; coefficient of, 512
Separator, cream, 85
Series resistance, law of, 447
Seven-in-one apparatus, 38
Shock, return, 389
Shunt, 450; ammeter, 488
Simple harmonic motion, 129–133; angular simple harmonic motion, 136

Sine, of angle, defined, 626
Sine law of refraction, 627; general law of refraction, 634
Siphon, 50
Size of image, law of, 615, 647
Sliding friction, 121; table of values of coefficient, 122
Slug, 72, 73
Snell's law of refraction, 627
Snow, 301
Soil, capillary action, 188; diffusion of water vapor, 192, 283; cooling due to evaporation, 285
Solidification. *See* Fusion
Solutions, freezing points, 273; boiling points, 283
Sound, 333–359; speed of, 335
Sounding, deep-sea, 339
Sound-ranging, 337
Spark at break of circuit, 510
Sparking potentials, table, 427
Specific gravity, defined, 24, 30; measurement of, 31–34; table of numerical values, 32; table of gases, 47
Specific heat, defined, 231; experimental determination, 232; numerical values, 233; of gases, 244
Specific inductive capacity. *See* Dielectric constant
Spectrum, of light, 634, 665–676, 691–694; of X-rays, 694–697
Speed, distinguished from velocity, 60; angular, 139–141; of water-waves, 330; of sound, 335, 354; of light, 600–605
Spherical aberration, 621, 648
Spheroidal state, 288
Standing waves, 331
Starting friction, 120
Stationary wave, 331
Steam engine, 317–322
Steam turbine, 321
Stefan's law, 263
Storage cell, 474–476
Strain defined, 196, 199
Strain theory of electric field, 398
Strength, ultimate, 201
Strength of magnetic field. *See* Field, magnetic
Stress, defined, 196; ultimate, 201
Strings, vibrations of, 350–353
Sun dog, 676
Surface tension, 182–189; table of numerical values, 186
Swing, 152
Symmer, Robert, 402

Telephone, receiver, 413; transmitter, 508; simple circuit, 508; use of thermionic amplifier, 566
Telescopes, 658–662
Temperature measurement, 203–208
Temperature scales, Fahrenheit, 203; centigrade, 203; Réaumur, 203; absolute, or Kelvin, 204, 222; gas, 204, 222, 225
Tension, surface, 182–189
Thales, 360, 401
Thermal capacity defined, 231
Thermionic amplifier, 565
Thermionic rectifier, 564
Thermodynamics, 246; first law, 246; second law, 246
Thermoelectric thermometer, 207
Thermometers, 204–207; maximum and minimum, 205; clinical, 206; resistance, 206; gas, 207, 224; thermoelectric, 207; wet-bulb and dry-bulb, 296
Thermos bottle, 261
Thunderstorm, 311
Tides, 160
Timbre, 343
Toggle joint, 8
Tone quality, 343, 352
Top, 143
Tornado, 152, 313
Torque, defined, 9; addition of, 10; on a magnet, 376
Torricelli's theorem, 172
Total reflection, 629
Tourmaline, polarization of light, 699
Transformer, 499
Transmission of electrical energy, 440, 518
Transmission lines, power loss, 440
Transmitter, telephone, 508
Transverse waves, 326
Triple point, 286
True ampere, 554
Tumbling in gymnasium, 153
Tungar rectifier, 564
Tuning-fork, mode of vibration, 353–354
Turbine, water, 173; steam, 321
Twinkling of stars, 637

Ultimate stress, 201
Ultra-violet light, 568, 673, 674

Undamped electrical oscillations, 585
Undercooling, 272
Units, fundamental, 16; derived, 17; arbitrary, 17; electrical, 550–553

Vacuum-tube amplifier, 565
Vacuum-tube generator, 586
Vacuum-tube rectifier, 564
Vapor, saturated, 275; nonsaturated, 277
Vaporization, heat of, 284. *See also* Evaporation
Vector, defined, 7; component of, 7; rotational, 142
Velocity, defining equation of, 54; uniformly accelerated, 54–64; as distinguished from speed, 60; of water-waves, 330; of sound, 335, 354; of light, 600–605
Ventilation, 255
Venus, mean temperature, 259
Vibratory motion, 129–137
Virtual image defined, 610
Viscosity, 125
Volt, unit of electromotive force and difference of potential, 408, 427, 552, 553
Volta, 402, 410, 465
Voltage, 418–421; of cells, 420; drop on wires, 424, 451
Voltaic cell, 407, 465; in series, 454; in parallel, 455; in series-parallel, 455; theory of, 471; gravity cell, 473; dry cell, 474
Voltameter, 471
Voltmeters, 419; electrostatic, 419; direct-current, 488
Volume elasticity, 198
Vowel sounds, 344

Washbowl, water flowing from, 152
Water, maximum density, 218
Water turbine, 173
Water vapor, pressure of, 275; numerical values of pressure, 277, 280, 296
Water-waves, 329
Water wheels, 173
Watson, Sir William, 403
Watt, James, 317
Watt, unit of power, defined, 102
Watt-hour, 430
Wave-length, defined, 326; of light, 262, 673, 683, 694; of X-rays, 697
Wave-motion, 325–333
Wave theory of light, 597, 625, 687, 707
Waves, heated bodies, 262; transverse, 326; longitudinal, 327; water, 329; sound, 334; electrical, 587–591; light, 598
Weather, 293–315
Weight distinguished from mass, 69, 73, 74
Wheatstone bridge, 446
Wheel and axle, 111
Whirlwind, 313
Wilcke, Johann, 402
Winds, 256, 304–314; effect of motion of earth on, 147, 306
Wireless, apparatus, 583–586; electrical waves, 588–591. *See also* Oscillations
Work, 90–106; defined, 91; units, 92, 430; table of numerical equivalents, 92, 239; general method of computing, 94; principle of, 109; electrical, 426, 428, 430

X-ray tubes, 562
X-rays, 561–563, 570, 578; ionization of gases by, 556; spectra of, 694; wave-lengths, 697

Yard, standard, 18
Young's modulus, 196; table of values, 197